Tax Me, But Spend Wisely
Public Finance and Government Accountability

Lucie Gadenne

University College London & Institute for Fiscal Studies *

November 2012

Abstract

This paper shows that local governments are more accountable when a larger fraction of their resources comes from local taxes. I construct a principal-agent model of public finance in which public revenues come from taxes and inter-governmental transfers. An increase in taxes changes the equilibrium allocation of public revenues towards more public goods and less political rents because citizens have better information on taxes than on transfers. I then compare how local governments in Brazil spend increases in tax and transfer revenues. Variations in tax revenues are created by a program that seeks to increase local tax capacity. Using a difference-in-differences methodology and quasi-exogenous variations in the timing of program take-up I find that the program increases tax collection of local governments by 11% after four years. I use several discontinuities in the rule allocating federal transfers to local governments to consider the impact of an exogenous increase in transfers. Results show, in line with the model's predictions, that an increase in local tax revenues leads to a bigger increase in local public services (health and education) than an increase in transfers of the same amount. Moreover extra transfer revenues lead to more corruption, extra tax revenues do not.

JEL classifications: D72, D73, H20, H71, H77

1 Introduction

There is a growing body of evidence that governments in developing countries waste or divert a substantial share of public revenues. Many economists and policy makers are consequently skeptical that making more funds available to these governments would lead to better development outcomes (Easterly (2008)). One common feature of the studies in this literature however is that they all consider how local governments spend non-tax revenues: Reinikka and Svensson (2005) find that schools in Uganda receive only a small share of funds allocated to them by the central government, Olken (2007) similarly finds that more than 20% of grants that local governments in Indonesia receive to finance road projects are diverted, whilst Svensson (2000) finds some evidence that aid increases corruption in politically divided countries. In Brazil, Caselli and Michaels (2011) and Ferraz and Monteiro (2010) show that windfalls from oil royalties lead to no improvement in local public good provision and Brollo *et al.* (2012) find that local governments that receive higher grants from the federal government become more corrupt.[1]

These studies are silent on whether their results would generalize for increases in these governments' tax revenues. There is however a long history to the idea of a strong and consistent connection between how governments are financed and how they spend their revenues, starting with Schumpeter (1918).[2]

In this paper I argue that the extent to which local governments are financed by tax revenues they collect as opposed to intra-governmental transfers affects the extent to which public revenues are diverted. The more they rely on tax collection, the more local politicians have to respond to their constituents' demands when allocating public spending and the less rents they can extract for their private use. I construct a theoretical framework that predicts that marginal increases in taxes will be more *accountability-inducing* than marginal increases in transfers based on standard political economy mechanisms. Evidence supporting this prediction is found by comparing the marginal propensity to spend from tax and transfer revenues of local governments (municipalities) in Brazil on municipal supply of public services and corruption.

The theoretical framework consists of a political agency model of public finance in which public revenues come from endogenous local taxes and exogenous transfers. A rent-seeking incumbent politician decides how to allocate the public budget between public good provision and diversion of funds for his private use (corruption). The key assumption is that tax revenues are perfectly observed by all players but transfers are a random variable whose realization is only fully observed by the politician. Information asymmetries lead to a difference in citizens' capacity to control the allocation of tax and transfer revenues. The model's key prediction is that an increase in tax revenues thanks to a decrease in collection costs shifts the allocation of public revenues towards more expenditures that benefit citizens, at the expense of corruption expenditures.

I test the model's predictions by comparing how local governments in Brazil (municipalities) spend increases in tax and non-tax revenues. I consider two policies that increase tax or transfer

---

[1]A more optimistic result is found in Litschig (2011) who finds that these same grants also lead to better education outcomes.

[2]This idea is central to interpretations of the emergence of representative governments and democracy in OECD countries. See for example North and Weingast (1989) for an account of how the economic elites in 17th Britain century obtained what they wanted (secure property rights and political representation) from the Stuart dynasty in exchange for larger taxation and borrowing opportunities.

revenues. The tax policy is a program that lowers municipalities' tax collection costs. Selection in the program is voluntary. The challenge to identification is thus that governments' choice to participate may not be orthogonal to unobservable factors that also affect tax collection and/or the allocation of public revenues. To study increases in transfers received by municipalities I consider variations in federal transfers municipalities receive that are created by a rule which specifies that transfers increase discontinuously in population size, replicating the identification strategy in Brollo *et al.* (2012), Litschig and Morrison (2010) and Litschig (2011).

My empirical strategy relies on a difference-in-differences estimator and a novel panel dataset on local governments' sources of revenue and local expenditure outcomes: the quality and quantity of municipally-funded supply of health and education and corruption of local politicians, obtained from randomized audits of local governments' public finances. A key characteristic of the tax policy is that municipalities decide when to apply to the program but the date at which they start one is determined by constraints faced by the supplier of the program. These create variations in the timing of program take-up that are unrelated to local characteristics. This specificity allows me to disentangle the impact of the program on tax revenues and public spending outcomes from that of (potential) time-varying determinants of selection that are unobserved.

Results show that the tax policy raises local tax revenues by 11% after four years and that this increase persists over time. The cost of the investments in tax administration are on average recovered after two years in the program. The rise in taxes is used to finance a 6% increase in municipal education infrastructure, an increase in school quality, and a 7% increase in municipal health infrastructure. Comparing the impact of this increase in taxes to that of an increase in transfers I find that an extra 10 Rs (5 USD) per capita of public revenues increases municipal health and education infrastructure significantly more when it comes from local taxes than when it comes from federal transfers. The share of corruption in total revenues increases when transfer revenues increase and decrease when tax revenues increase, as predicted by the model.

The impacts of taxes and of transfers on public spending outcomes are estimated on different sub-populations of Brazilian municipalities. The observed differences between the marginal propensity to spend on taxes and transfers could therefore be due to differences in propensities to spend out of any type of revenue between these populations. The strategy used to identify the impact of transfers allows me to rule out this explanation with reasonable confidence. The transfer policy creates 15 population thresholds at which the amount of transfers jumps discontinuously. I therefore estimate 15 local treatment effects on different sub-population of municipalities, some of which look extremely similar to the municipalities that increase their taxes thanks to the program. These estimates provide plausible bounds for the underlying heterogeneity of the effect of transfers on public expenditure outcomes in the population.

To the best of my knowledge this paper is the first to considers the effect of increases in tax revenues on public spending outcomes and to compare how two different sources of revenues are spent by the same government units. Several authors argue that there is a causal relation between the extent to which governments are financed by taxation and how accountable their are (see in particular Moore (2007) and Brautigam *et al.* (2008)). Evidence in line with this idea has been found, based on case studies (Gervasoni, 2010), correlations across countries (Ross, 2004) or lab in the field experiments (Paler, 2012). Closely related to this paper, Fisman and Gatti (2002) establish a positive relationship between the proportion of US states' revenues derived from federal

transfers and the number of convictions of public employees for abuse of public office.[3] Similarly Zhuravskaya (2000) provides evidence that Russian cities that keep more of the tax revenues they collect provide more public goods.

This paper builds on these previous results in several ways. First, it relies on using variations in tax and transfer revenues that stem from clearly identified sources. Second, the richness of the data allows me to consider the marginal propensity to spend out of tax and transfer revenues on a large range of municipal expenditure outcomes. Third, it offers a simple theoretical mechanism to explain why taxes and transfers are spent differently and some evidence that this mechanism plays a role in the Brazilian context, though other mechanisms may also be at play.

The theoretical underpinnings for the idea that relying on local taxes affects political officials' incentives date back to at least Tiebout (1956). The more recent literature on market-preserving fiscal federalism argues that the more politicians depend on locally generated revenue the more they will invest in public goods that increase their local tax base (see Weingast (2009)). This paper's theoretical framework differs by relying on an explicit political economy mechanism to explain why taxes lead to more spending on public goods and less corruption than transfers. The mechanism outlined here will hold even if local governments cannot finance growth-enhancing local public goods and taxpayers are not mobile.[4] It is also closely related to previous political agency models which argue that information asymmetries lead to more rent-taking opportunities by politicians (see Besley (2006) for a review). Those do not however explore the possibility that public revenues are more or less well observed depending on their source and that this will affect elected officials' accountability to their constituents. Another related literature is that on the fly-paper effect, that seeks to explain why local governments spend more from intergovernmental grants than from equivalent increases in local private income. Several authors in this literature point out that the fly-paper effect follows naturally from the assumption that citizens have imperfect information about the level of grants (Strumpf (1998))- a simple extension of the theoretical framework developed here would indeed predict the fly-paper effect.

By estimating the returns to investments in tax capacity this paper finally contributes to the growing literature on state capacity and development. This literature argues that governments' investment in their capacity to tax constitutes an important covariate of economic development (Besley and Persson, 2009, 2010, 2011, Cagé and Gadenne, 2012). Widely optimistic prognoses regarding the efficiency of investments in developing countries' tax administrations abound: the President of the African Tax Administration Forum, Oupa Magashula, for example claimed that investing in public resource mobilization can have up to '*a tenfold multiplier effect on states' resources*'(OECD, 2010b). Recent papers have studied policies that seek to increase tax collection (see Pomeranz (2010), de Paula and Scheinkman (2010) and Monteiro and Assuno (2011)), however none estimates their impact on tax collection. I find an annual multiplier effect of just over one, far from tenfold but still very cost-effective.

The outline of the paper is as follows. Section 2 provides an agency model of public finance

---

[3]As the authors point out the causal interpretation of this relationship is limited by potential endogeneity problems. These problems are mitigated in this paper by the use of variations within municipalities over time and corruption indexes from randomized audits.

[4]The type of public good provision this paper considers is unlikely to have the type of short-run growth effect required for the mechanism outlined in Weingast (2009) and Zhuravskaya (2000) to be relevant amongst Brazilian local governments, as I explain below.

3

that relates how governments are financed to how they allocate their budget between private rents and public good provision. Section 3 presents the institutional context and the tax and transfer policies. Section 4 shows that these two policies generate large variations in tax and transfer revenues and Section 5 compares the marginal propensity to spend from tax and transfer revenues on local public services and corruption.

2 Model

2.1 Set-Up

This model follows the political agency framework of Besley and Smart (2007) in which a representative citizen decides whether to re-elect a rent-seeking incumbent without observing part of his actions. The budget of the government is representative of that of local governments throughout the world: public resources R come from local taxes T, endogenously determined, and intergovernmental-transfers F which are exogenous and subject to some random variation. Transfers can take two values : F is equal to $F_H = \bar{F}(1+u)$ in the high state H with probability q and $F_L = \bar{F}(1-u)$ in the low state L, where $u, q \in [0,1]^5$.

The incumbent politician faces a budget constraint $T + F = R = G + S$, with G the level of public good and S the rents he diverts for himself. He maximizes the sum of rents extracted from being in office $S + \sigma Z$, where Z is the exogenous value of re-election and σ the probability of re-election. He can choose to divert all public resources and forgoe re-election but institutional constraints limit maximal rent taking to $\bar{S} = \alpha R$ where $\alpha < 1$.[6] Challengers in the election would behave in the same way as the incumbent once elected; the election is a way for the citizen to discipline the incumbent, not to choose the best type of candidate.

The representative citizen derives utility from the provision of public good net of taxes. Her welfare is $W(G,T) = G - \phi C(T)$ where ϕ indexes the marginal utility cost to the citizen of paying taxes and $C(.)$ is increasing and strictly convex. I define $h(\cdot) = C'^{-1}(\cdot)$.

2.2 Full information equilibrium

The citizen chooses for each state $i = H, L$ the reelection rule $\sigma(G_i, T_i) = \sigma_i$ that induces the politician to provide the policy menu (G_i, T_i) that maximizes her welfare. The maximum level of public good G_i she can obtain from the government when paying taxes T_i must be so that it leaves the government with enough rents today to make abiding by the re-election rule more attractive than running away with maximum rents and forgoing re-election. This *participation constraint* takes the form:

$$T_i + F_i - G_i + \sigma_i Z \geq \alpha(T_i + F_i), \forall i = H, L \tag{1}$$

---

[5]One can alternatively think of F as any source of public revenues that is not directly extracted from citizens, such as revenues coming from the government's sale of natural resources, profits of public monopolies or development aid. The predictions of the model are thus also relevant at the level of federal government.

[6]I assume $Z < \alpha R$ to ensure that rents are never negative. This assumption simply says that the politician discounts the future and cannot expect to extract more rents in the future than the maximal level of rents it could extract today.

Re-electing the incumbent leads to an increase in the public good at no cost to the citizen so in equilibrium she sets $\sigma_i^* = 1$ in each state i as long as the government provides the menu (G_i^*, T_i^*) such that:

$$G_i^* = (T_i^* + F_i)(1 - \alpha) + Z. \tag{2}$$

T_i^* is set such that the marginal value of the public good is equal to the marginal cost of taxation: $T_i^* = h(\frac{1-\alpha}{\phi})$. Local taxes are decreasing in the marginal cost of paying taxes ϕ and in α, a proxy for the ease with which the politician can run away with public resources.

When the citizen fully observes all public revenues the way in which in the local government is financed does not matter. The marginal effect of an increase in taxes or transfers is to increase the public good by $(1 - \alpha)$ and rents by α.

2.3 Equilibrium with asymmetric information

Assume now that the citizen does not perfectly observe transfer revenues: the realized value of F is known only to the incumbent. The citizen perfectly observes the taxes she pays. Asymmetries of information increase the incumbent's capacity to extract rents from the public budget as he can now pretend to be in the low state when he receives high transfer revenues to capture the difference in revenues between the high and the low states. A formal proof of this result is given in the paper's theoretical appendix, I sketch the intuition below.

To deter the incumbent in state H from implementing the L state menu the menus offered by the citizen must now also respect the incentive constraint:

$$S_H + \sigma_H Z = T_H + \bar{F}(1 + u) - G_H + \sigma_H Z \geq T_L + \bar{F}(1 + u) - G_L + \sigma_L Z \tag{3}$$

And similarly for the incumbent in state L:

$$T_L + \bar{F}(1 - u) - G_L + \sigma_L Z \geq T_H + \bar{F}(1 - u) - G_H + \sigma_H Z \tag{4}$$

Putting together (3) and(4) there is only one situation in which both constraints are satisfied simultaneously : $G_H = G_L + T_H - T_L + Z(\sigma_H - \sigma_L)$.

Intuitively it is still optimal for the citizen to ask the incumbent in the low state to provide the maximal amount of public good given the amount of taxes paid: state L's *participation constraint* – equation (1) – is binding (see Appendix for a formal proof). This implies the following equilibrium levels of public good provision:

$$G_L^* = (T_L^* + \bar{F}(1 - u))(1 - \alpha) + \sigma_L^* \ Z \tag{5}$$

and

$$G_H^* = (T_H^* + \bar{F}((1 - u))(1 - \alpha) + \sigma_H^* Z + \alpha(T_H^* - T_L^*) \tag{6}$$

Re-election leads to an increase in the public good at no cost to the citizen whatever the state, so $\sigma_H^* = \sigma_L^* = 1$. Maximizing $W(G_H, G_L, T_H, T_L; q)$ subject to (5) and (6) determines the level of taxation in both states :

$$T_H^* = h(1/\phi) \tag{7}$$

and

$$T_L^* = \max\{0; h((1 - q - \alpha)/\phi(1 - q))\} \tag{8}$$

5

It is optimal for the citizen to pay less taxes in the low state as any increase in the level of taxes offered in the low state menu makes mimicking the low state equilibrium more attractive to the incumbent in the high state. This comes at the cost of less public good in the low state. The less likely the low state (the higher q) the more the citizen is willing to incur this cost, and the lower T_L^*. The asymmetry of information leads to an equilibrium with lower public good provision (on average) than in the full information equilibrium due to the increase in rent-seeking obtained by the incumbent in state H.

The structure of public finance now affects the way in which the incumbent allocates the budget. Using equations (5) and (6) we can write the average level of the public good as:

$$E(G^*) = (1 - \alpha)(E(T^*) + \bar{F})) - \bar{F}u(1 - \alpha) + Z + (1 - q)\alpha(T_H^* - T_L^*) \tag{9}$$

The term $u(1-\alpha)\bar{F}$ corresponds to the informational rents the incumbent can appropriate in state H by 'hiding away' the extra transfer revenues. The last term in equation (9) simply says that the more the citizen can provide the incumbent in the high state with high powered incentives relative to the low state (the bigger the difference between taxes in both states) the lower the informational rents. A marginal increase in taxes still increases public good provision by $(1 - \alpha)$, assuming for simplicity that the increase does not affect the spread $T_H^* - T_L^*$. A marginal increase in average transfers has a smaller impact of $(1 - \alpha)(1 - u)$.[7] Finally, note that the higher the asymmetry of information (higher u) the bigger the difference between the marginal impact of taxes and transfers. At the limit when $u = 1$ any increase in transfers is spent fully on rents, and when $u = 0$ the equilibrium is a full information one.

The equilibrium share of rents in public revenues s^* is increasing in the share of transfers in the budget proxied by $\bar{f}^* = \bar{F}/E(R)$:

$$E(s^*) = \alpha + E(\bar{f}^*)2u(1 - \alpha)(1 - q) - Z/E(R) - (1 - q)\alpha(T_H^* - T_L^*)/E(R) \tag{10}$$

Equations (9) and (10) summarize the *accountability effect of taxes on the allocation of public spending*: they show that as the share of taxes in revenue increases, so does the share of revenues that is spent towards public good provisions. Intuitively increasing the share of taxes increases the amount of information the citizen has on her government's budget and thus limits the extent to which a rent-seeking politician can capture public funds by 'hiding' them. This leads to an allocation of the budget that is more favorable to the citizen.

2.4   Impact of a tax capacity program

Consider now the impact of a program that makes the tax administration more efficient. This takes the form of a smaller difference between the cost of taxation borne by taxpayers $\phi C(T)$ and how much taxes go in the government budget T: the program decreases ϕ.[8] This makes the

---

[7] I assume throughout that any increase in transfers *ceteris paribus* comes from an increase in $\bar{F}$ and not a change in the probability q of the high state. This is consistent with the type of increase in transfers considered in the empirical strategy which are a consequence of a local government moving to a higher transfer bracket, not a random shock to transfers.

[8] One could also model the efficiency of the tax administration by introducing a cost to the government of levying taxes. The reform would then lower that cost, leaving the results of this model unaffected. I explain in this paper's online Appendix that the program I consider does decrease the cost of paying taxes.

citizen more willing to pay taxes in order to get more public good. Using equations (7) and (8) the impact of a program that lowers the efficiency cost by $d\phi < 0$ on taxes is given by :

$$\frac{\partial E(T^*)}{\partial \phi} d\phi > 0 \tag{11}$$

The program will also lead to an increase in public good provision proportional to the increase in taxes :

$$\frac{\partial E(G^*)}{\partial \phi} d\phi = (1 - \alpha) \frac{\partial E(T^*)}{\partial \phi} d\phi + (1 - q) \alpha \frac{\partial (T_H^* - T_L^*)}{\partial \phi} d\phi > 0. \tag{12}$$

Because it decreases the share of transfers in total revenues f^* the reform also lowers the share of rents s^* (equation (10)). Comparing the marginal propensity to spend from extra taxes thanks to the program or from an increase in the average value of transfer revenues($E(F)$) yields two testable propositions:

Proposition 1 *The rise in taxes due to the reform leads to more increase in public good provision than a rise in transfer revenues of the same amount $\frac{\partial G^*}{\partial E(T^*)} > \frac{\partial G^*}{\partial E(F)}$.*

Proposition 2 *The rise in taxes due to the reform leads to a fall in the share of rents in public revenues. An increase in transfer revenues increases the share of rents in public revenues : $\frac{\partial s^*}{\partial E(T^*)} < 0 < \frac{\partial s^*}{\partial E(F)}$.*

A final proposition comes from observing that an increase in the information the citizen has on the budget lowers the equilibrium information rents and thus mitigates the relative accountability effect of taxes relative to transfer revenues:

Proposition 3 *The higher the information the citizen has on the level of transfers (the lower the u) the more similar the impact of an increase in taxes thanks to the program and the impact of an equivalent increase in transfers: $\frac{\partial G^*}{\partial E(T^*)} - \frac{\partial G^*}{\partial E(F)}$ and $\frac{\partial s^*}{\partial E(T^*)} - \frac{\partial s^*}{\partial E(F)}$ are increasing in u.*

Formal proofs of these propositions and all the results in this section are in the theoretical appendix.

This simple sketch shows that if we assume citizens do not perfectly observe transfer revenues increases in taxes are spent 'better' (more on public good provision, less on rents) than increases in transfers. The assumption that citizens perfectly observe tax revenues, though natural in a model in which taxes are endogenous to the contract between a representative citizen and a politician, may be too strong in practice. One can also make the weaker assumption that asymmetries of information are less strong for tax than for transfer revenues: citizens are likely to observe increases in taxes better than increases in transfers, even in they are not perfectly aware of the level of total taxes. In the theoretical Appendix I show that the above propositions still hold if we assume that citizens only observe part of tax and transfer revenues but have slightly better information on taxes.

3 Context : Brazilian local governments

3.1 Local public finances in Brazil

Tax revenues and the tax policy

The 1988 Brazilian constitution devolves substantial expenditure responsibility and tax autonomy to the country's more than 5500 local governments.[9] The rates and bases of three main local taxes (a service tax, a property tax and a property sales tax) as well as the method of tax assessment and collection are decided by local elected officials.

Municipalities' *de facto* tax collection is small. They collect less than 15% of their total revenue themselves (roughly 2% of GDP). The spiralling of local debts in the early 1990s has directed much policy attention in Brazil towards the low tax efforts of local governments with commentators pointing out the poor quality of local tax administrations (see for example Afonso (2005)). The few studies of Brazilian tax administrations available paint a picture of unskilled and overworked staff with outdated tax registers, no institutional memory and a lack of methods to accurately assess tax liabilities.[10] High costs of understanding and paying taxes likely push many citizens into non-compliance and in the early 2000s some local officials publicly admitted to tolerating a situation of ongoing tax amnesty where tax arrears are rarely recovered (Afonso and Araujo (2006), BNDES (2002)).

The *Programa de Modernizacao de Administracao Tributaria* (*PMAT*) was launched in 1998 by the Brazilian Development Bank (*BNDES*) to increase municipalities' capacity to tax their citizens. It provides all local governments that apply with subsidized loans to invest in modernizing their tax administration. 331 municipalities started a program between 1998 and 2008, covering 40% of the Brazilian population.

The program's loans can only be used to fund investment expenses related to the tax administration, other budget items are explicitly not eligible. *BNDES* staff check the receipts for all expenditures made in relation to the program but otherwise exert no control on the public finance processes of participating municipalities. Fieldwork suggests municipalities used program loans to create or update tax registers, decrease the costs of paying taxes by increasing the number of tax offices, means and frequency of payments, and facilitate controls of tax payers through the recovery of tax arrears or the development of cross-checking mechanisms. The paper's online Appendix describes the context of the program's creation and the type of investments in tax capacity it funded in more details (see also Santos *et al.* (2008), BNDES (2002) and Afonso and Serra (1999)).Table 2 shows that in 2003 municipalities that had already started a *PMAT* program were much more likely to have updated their property tax register since 1998 than both the average municipality and municipalities that joined *PMAT* since.

The timing of application to and entry in the *PMAT* program is of particular interest. Municipalities typically apply to the program by submitting a first tax modernization project, then wait between one and four years before they actually start the program, ie receive their first loan. Table 14 shows that the time between applying and starting a program varied over the period.

---

[9]Bardhan and Mookherjee (2006) classify Brazil as one of the few developing countries in which local governments have been given substantial tax autonomy.

[10]An extensive study of the property tax collection in Brazil's largest metropolitan areas estimates that less than 60% of urban property is registered on any tax administration's files (de Carvalho Jr. (2006)).

Table 1: Municipalities that updated their tax registers by program status

	% Updated 1998-2003	Nb municipalities
Never joined the program	72%	4723
Started after 2003	70%	122
Started before 2003	85%	146

Source: *Perfil dos Municipios Brasileiros, 1998, 2004.* % of municipalities which have updated their property tax registers between 1998 and 2003 in the first column (% of those that respond to the question), number of municipalities which updated their property tax registers during that period in the second column.

This is due to variations in the supply of the program. The *BNDES* processed all applications itself for the first three years of the program's existence, taking more than 2.5 years per application on average. It contracted the public bank *Banco do Brasil* in 2002 to take in charge most of the application process.[11] *Banco do Brasil*'s involvement initially accelerated the application process until the bank decided to cut down the resources allocated to *PMAT* in 2006. It administered most of the projects from 2002 to 2005, and substantially shortened the waiting period. In 2007 the *BNDES* signed a similar agreement with another public bank, the *Caixa General* and we see another decrease in waiting time for the 2007 cohort.

Table 2: Average time between program application and program start

Application Year	Years to Program Start	Nb Municipalities
1997	2.6	11
1998	2.7	8
1999	2.5	18
2000	2.3	21
2001	1.5	41
2002	1.1	60
2003	1	55
2004	0.7	40
2005	1.4	20
2006	1.8	36
2007	0.7	22
All	1.2	331

Source: *BNDES* data collected by the author.

The possibility that the most eager municipalities pressure the *BNDES* to start the program soon after applying cannot be ruled out. However, in the overwhelming majority of cases (95%) the order in which municipalities apply to the program corresponds to the order in which they start a program. This suggests that whilst municipalities choose when to apply the precise date

---

[11]The *BNDES* is based in Rio de Janeiro but the *Banco do Brasil* has branches around the country, allowing for more geographic outreach.The *BNDES*' own devotion of resources to the program varied over the years: the federal government's initial push for the policy was short-lived and in 2001-2002 only one *BNDES* official was working on *PMAT*, the idea being that *Banco do Brasil* would take charge of most of the administrative work. The swearing into office of a new President in 2003 put the project back up high on the *BNDES* agenda and today the staff team has stabilized to around 12 individuals.

at which they start a program is out of their control: we see in Figure 14 that the distribution of application dates over time is relatively smooth while that of start dates displays clear bunching for the years 2002-2004. It also provides, for each municipalities, two dates of interest for identifying the program's impact which I return to below.

Figure 1: Distribution of Application and Start Dates
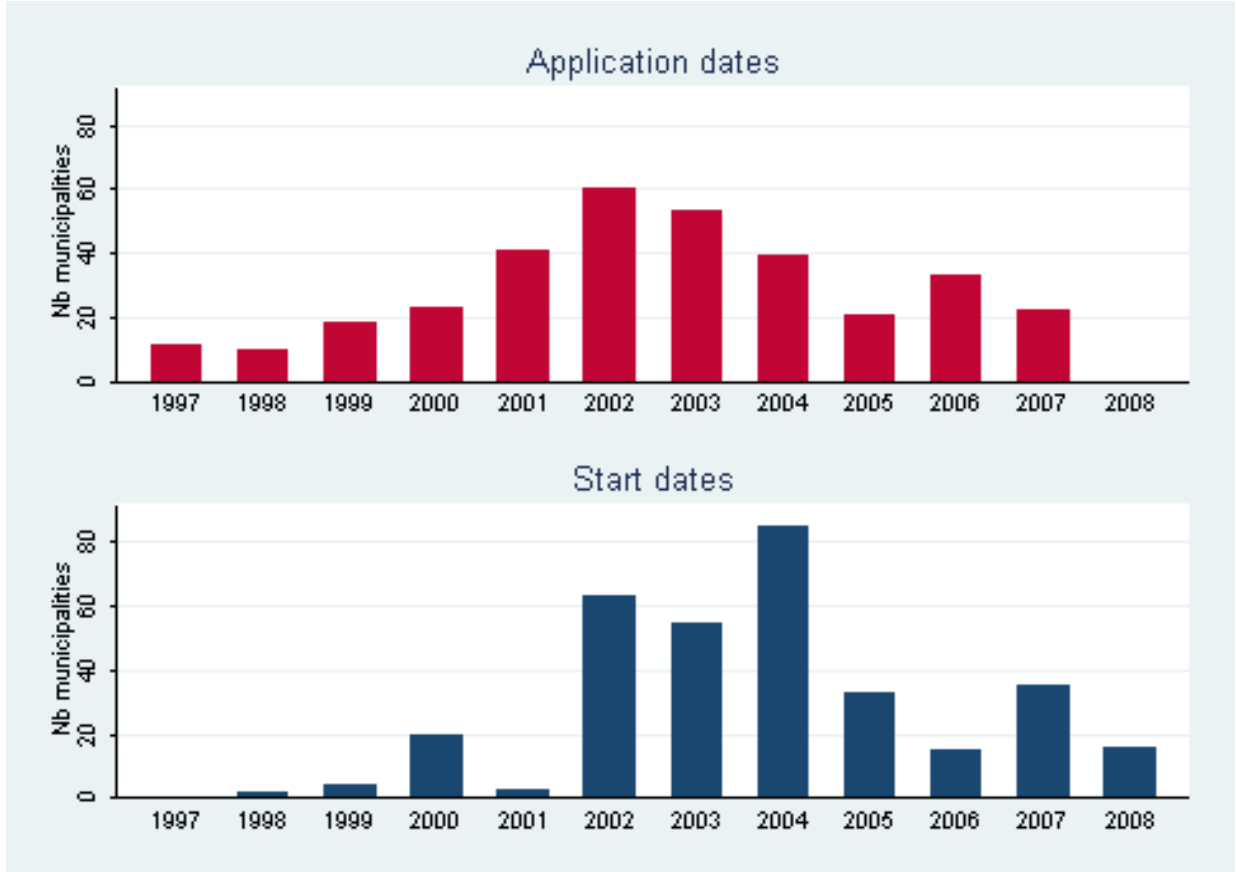
The transfer policy

Local revenues which are not locally levied come from inter-governmental transfers. The largest transfer is the *Fundo de Participacao dos Municipios* (FPM) federal transfer which is constitutionally mandated and the largest source of local revenues (40%). I focus on this transfer for two reasons. First, as outlined below, its allocation depends on population size via a formula based on cutoffs which is crucial for my identification strategy. This feature has attracted the attention of researchers that seek to identify the impact of higher transfers on various local outcomes (see Brollo *et al.* (2012), Litschig and Morrison (2010) and Litschig (2011)), using it as a source of identification is not a novelty here. What is novel to this study is the comparison of how local tax revenues and transfers are spent. The key advantage of considering *FPM* transfers to do

such a comparison is the fact that local governments are free to spend these transfers as they see fit, whereas other transfers are earmarked for certain types of expenditures. Local politicians therefore have the same discretion in deciding how to spend tax and *FPM* revenues.

The *FPM* allocation rule specifies that all municipalities in the same state and in a given population bracket receive the same amount of transfers. The revenue sharing mechanism determining the amount $FPM_{i,t}^s$ received by government i in state s is

$$FPM_{i,t}^s = \frac{f(pop_{i,t})}{\sum_{j \in s} f(pop_{j,t})} FPM^s \tag{13}$$

where $f(pop_{i,t})$ is the coefficient corresponding to the population bracket in which the local government's population is found and FPM^s is the total fund allocated to the state. Municipalities with more than 142,633 inhabitants (3% of the sample) receive an amount of FPM transfers based on a slightly different rule and are therefore excluded from the analysis of the impact of the transfer policy.

The coefficient of each municipality is set annually by the Federal Court (*Tribunal de Contas União, TCU*) based on the population estimates calculated by the Brazilian Institute of Statistics (*Instituto Brasileiro de Geografia e Estatística, IBGE*). *IBGE* uses a top-down approach for estimation which ensures that municipality estimates are consistent with state estimates, based on demographic evolution between subsequent Census. The Appendix describes the procedure for estimating local population and presents the population brackets and associated coefficients, both set by decree in 1981, in Table 9. Figure 2 displays the scatterplot of received and theoretical *FPM* transfers as a function of population size in the state of Sao Paulo in 2005 and 2009 to illustrate the rule (theoretical transfers are constructed by applying the rule on the population estimates the vertical lines represent the population thresholds). We see that, although there are multiple cases of mis-assignments around the population thresholds, the amount of *FPM* transfers received by municipal governments displays clear jumps at each threshold.

Data on participation to the *PMAT* program, date of application, program start, and amount borrowed through the program have been collected by the author at the *BNDES*. Public finance data for the years 1999-2009 from the *FINBRA* dataset documents annual tax collection, and official data on *FPM* transfers comes from the *Tesouro Nacional*. All revenue variables used in the analysis are per capita and in 2000 Reais.[12]

3.2 Local expenditure responsibilities

Brazilian municipalities are responsible for 17% of total public expenditures in Brazil. I focus on two types local public expenditure outcomes: the provision of municipal health and education to proxy for public good provision (G in the model above), and diversion of public revenues by the administration which proxies for the share of rents in public revenues (s).

The Brazilian constitution stipulates that states and municipal governments share the responsibility for the provision of primary and secondary education. In practice state governments manage secondary schools and municipal governments are mostly in charge of primary schools

---

[12]Per capita variables are computed using annual population estimates provided by the Brazilian statistical institute *IBGE* and deflated using the GDP deflator.

Figure 2: The transfer allocation rule
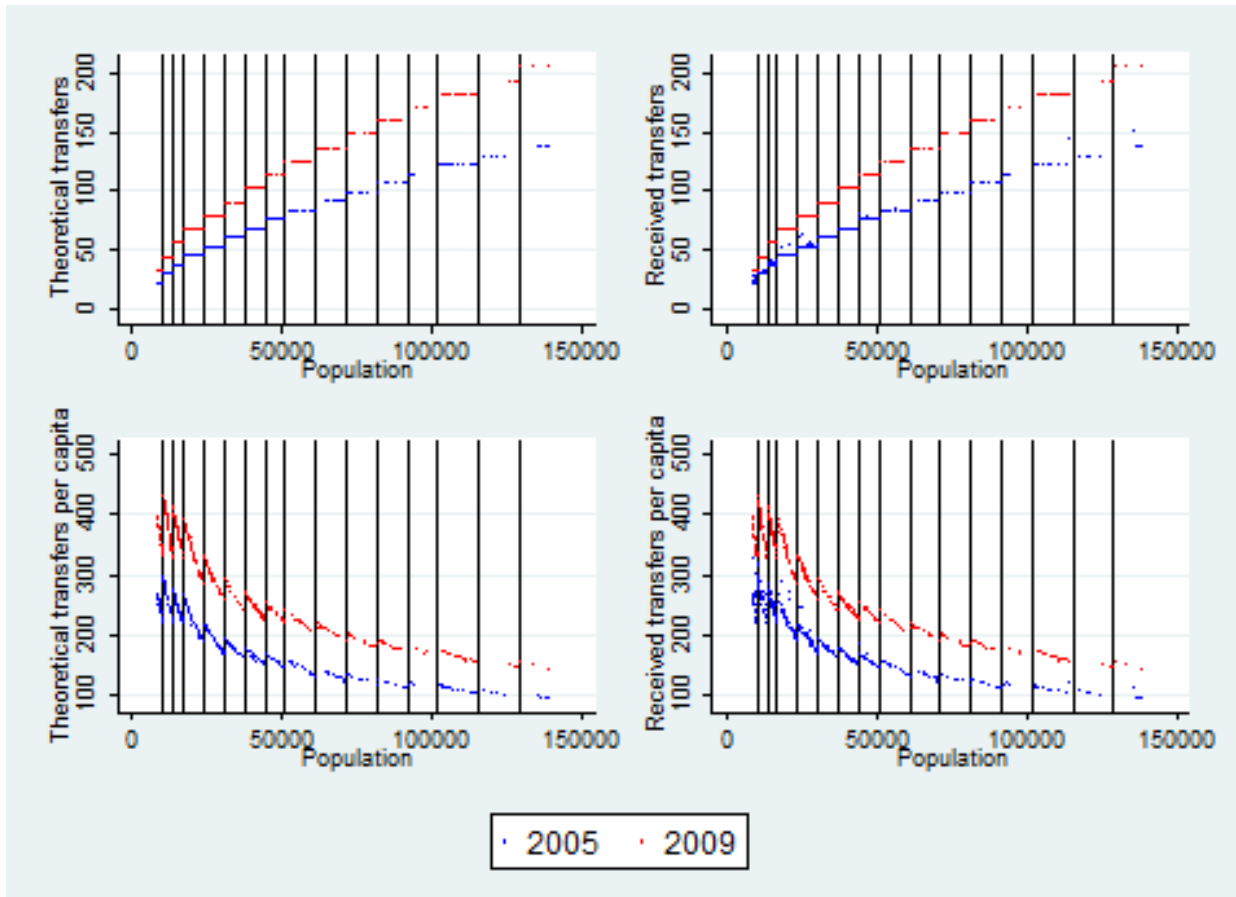
Graphs on the left present the amount of transfers the FPM allocation rule predicts each municipality will get in year t as a function of its estimated population size released on the 1st of July by the IBGE in the year $t-1$. Graphs on the right present the amount of transfers effectively received by the municipality as declared by the *Tesouro Nacional*. The top panel shows nominal total transfers, the bottom panel nominal transfer per capita. The blue dots (the lowest series) are for the year 2005, the red dots (the highest series) for the year 2009.

(*ensino fundamental*).In 2005 approximately 85% of all grade 1 to 4 schools were run by local governments, the remainder being private or state primary schools (see Ferraz *et al.* (2012)). Local governments provide infrastructure, school lunches and transportation and hire and pay teachers.

I use panel data on municipal education inputs from the annual census (*Censos Escolar*) of all Brazilian schools to measure the quality and quantity of municipal education infrastructure. To measure quantity I consider the number of classrooms in use in municipal schools per thousand inhabitants. Local governments receive substantial federal transfers directed towards education expenditure but those generally come with rules specifying that they must be spent on staff, school lunches or school transport and not on physical teaching infrastructure.[13] Classrooms are therefore the education input that is the most likely to be under-funded and that municipalities can most easily adjust in the short-run by furbishing existing unused rooms or renting extra space. Several variables are available to proxy for the quality of municipal education infrastructure: the number of schools with computers, with internet, with a sports facility, a library, television/video equipment and connected to the sewage and electricity systems. I use principal components analysis (PCA) to combine these seven measures into an index of infrastructure quality. The first principal component explains 60% of the variation in the data; this suggests that using PCA reduces the dimensionality of the data with little loss of information.

Education is the largest budget item of local governments, representing roughly a third of expenditures; health is the second largest budget item. Most of the responsibility for administering basic health care is in the hands of local governments through the Family Health Program (*Programa Saude de Familia*). Municipalities are in charge of hiring, paying and supervising medical teams and providing the infrastructure for the primary and preventive health units (Rocha and Soares, 2010). Data on municipal health infrastructure comes from the *Pesquisa Assistência Médico Sanitária*, a comprehensive survey of the Brazilian health system carried out by the Ministry of Health and available in 1999, 2002, 2005 and 2009. It documents the number of private, federal, state and municipal health units in each municipality. I use the number of municipal health establishments per 100,000 inhabitants to measure the quantity of municipal health infrastructure.

There is finally considerable information on how local governments divert public resources away from public uses in Brazil thanks to a federal anti-corruption program. Since 2003 over 1800 local governments have been randomly chosen by lottery to be audited by staff of the independent audit agency *Corregedoria Geral da União* (*CGU*). These staff audit the use of local revenues by municipalities over the two years prior to the audit by collecting administrative documents and interviewing citizens and administrative staff. They check for example whether spending can be accounted for by receipts, whether program rules are met, and whether procurement of public works is done competitively. The results of those audits are publicly available records. Ferraz and Finan (2011) estimate using this data that approximately 550 million US dollars per year were diverted in the period 2001-2003, or 8% of audited transfer revenues.[14]

The *CGU* audits municipal programs that are funded by discretionary transfers, there is no

---

[13]For example 60% of the largest of those education transfers, *FUNDEB*, must fund teacher's salaries. The *PNAE* transfer funds school lunches. All my results are unaffected when I control for the amount of education-specific transfers received by the municipality.

[14]For more on the anti-corruption program and analysis using data from the audits see Ferraz and Finan (2008a), Ferraz and Finan (2011), Ferraz *et al.* (2012),Litschig and Zamboni (2008) and Brollo *et al.* (2012).

direct audit of how the two types of local revenues considered in this paper- taxes and *FPM* transfers - are used. However most discretionary federal transfers require that municipalities contribute some of their 'own revenues' (constitutionally defined as *FPM* transfers or taxes) to the programs they fund so we can think of the audits as reflecting the overall quality of government spending (see Brollo *et al.* (2012)). Importantly for this study there is no reason to think that the use of tax revenues is more closely audited than that of *FPM* transfers, or vice versa.

The corruption data I use comes from the coding of the *CGU* audits for the years 2003-2006 used by Litschig and Zamboni (2008). It is available for a small sample of 971 municipalities, 54 of which join the program. Following the existing literature I construct a corruption index from this data by scaling the number of irregularities by the number of civil servants in the local government administration[15] and the total amount of government revenue audited. This provides a proxy for the share of diverted revenues in the total public budget, the theoretical outcome considered in the model above.

NOTE TO THE READER. My current analysis of corruption is limited by the small number of years available in the data, in particular my analysis of pre-treatment trends. I am expecting to obtain data on audits for the period 2003-2011 soon, this will allow me to perform more checks for the corruption outcomes. Results may also change when I use the bigger sample.

---

[15]This is obtained from the dataset *Perfil dos Municipios Brasileiros* 1998, published by the *IBGE*.

4 Impact of the tax and transfer policies on public revenues

To estimate the impact of increases in tax and transfer revenues on local public expenditure outcomes one must observe variations in tax and transfer revenues that are credibly exogenous to other determinants of expenditure outcomes. One would ideally like to randomly assign increases in tax or transfer revenues to some municipalities and compare their average outcomes. Policies that distribute substantial amounts of revenues across governments are hardly ever the subject of randomized trials so I turn to non-experimental methods that create a credible counterfactual under a reasonable set of assumptions. This section shows that the tax policy (the *PMAT* program) and the transfer policy (the discontinuity in the transfer allocation rule) create significant variations in tax and transfer revenues. The following section considers the impact of these variations on public expenditure outcomes.

4.1 The tax policy

Why did municipalities join the program?

The biggest identification concern is that municipalities that self-select in the *PMAT* program could be different from those that don't along dimensions which correlate with outcomes. I start by studying determinants of program participation to get a feeling of how large this concern is in practice.

Given the large number of municipalities that did not join the program I choose to exclude from my analysis those which field interviews and inspection of summary statistics suggest have no interest in the program and constitute a very poor counterfactual for the evolution of outcomes in treated municipalities. Those municipalities are those whose population or GDP are below the minimum value of those variables amongst the sample of treated municipalities. More specifically I take out municipalities with a population of less than 3,500 (10% of municipalities which did not join the program) or income per capita below 790 Rs (9%). Over 3,000 control municipalities remain to be included in the analysis.[16]

Table 3 presents results obtained by estimating a discrete time hazard model of the probability that a given municipality at a given period of time applies to the program as a function of both pre-treatment characteristics of municipalities and time-varying covariates.[17] It shows that municipalities that join the program are richer and more populated than the average Brazilian municipality. Program participants often say they joined the program because they were dissatisfied with their current level of tax collection compared to what they thought was their tax potential, we see that all else equal they indeed collect less taxes in 1998. They are also more educated,

---

[16]I present results obtained using all municipalities that never join the program as a control group as a robustness check below.

[17]See Jenkins (1995) for a description of the method and Galiani *et al.* (2005) for a similar application to privatization of local water provision in Argentina. Observations corresponding to municipalities which have applied to the program at least a year ago are dropped from the sample. The dependent variable is a dummy equal to 0 for municipalities which have not applied to the program yet and 1 the year in which they apply. Municipalities which joined in the first two years and observations for 1999 are excluded because the variable 'distance to the 5 closest municipalities which have already joined a *PMAT* program' is not available for those. Results are very similar when that variable is taken out of the specification and the sample is not restricted.

less agricultural and slightly more politically competitive – all these characteristics are highly serially correlated so the identifying variation used for these estimates is mostly cross-sectional. Political characteristics of the mayor do not play a role (a full set of 26 dummies for political parties do not come out as jointly or individually significant) and neither does alignment with the governor's party. This provides some reassurance that the program's loans were not directed towards politically favored mayors. Finally, field interviews suggest that part of the selection in the program may have been determined by which municipalities hear about it. There was very little advertising for the program and most participants said they heard about it once one of the municipalities in their neighborhood joined. I find that local governments whose neighbours have already joined a $PMAT$ program (proxied by the average distance between a municipality and its 5 closest $PMAT$ neighbours) are indeed more likely to join.

In the second column I consider whether past shocks determined program uptake, and find no evidence of an 'Ashenfelter dip' in tax revenues or that selection in the program is driven by specific economic, demographic or political shocks. Results in the third column suggest that treated municipalities followed similar trends to the control ones in the 1996-1999 period. The fact that no observable shocks determine selection in the program motivates the use of the difference-in-differences methodology described in the next section.

Empirical strategy

The determinants of selection in the program suggest that many of the unobservable characteristics that may confound identification are likely to be fixed over time; I estimate a difference-in-differences model to control for such time-invariant unobserved heterogeneity.

Formally, I estimate the model :

$$Y_{i,t} = \beta P_{i,t} + \delta X_{i,t} + \gamma_t + \mu_i + \epsilon_{i,t} \tag{14}$$

where $Y_{i,t}$ is tax collected per capita in municipality i in year t, $P_{i,t}$ is a dummy equal to 1 if municipality i started a program in a year $s \leq t$ and γ_t and μ_i a set of year and municipality fixed effect. Time-varying controls in $X_{i,t}$ are GDP per capita, population, and political characteristics of the municipality (competitiveness of the last election, political party of the mayor, term limit). All specifications below allow for arbitrary covariance structure within municipalities by computing standard errors clustered at the municipality level.[18]

The key identifying assumption required for the interpretation of β in (14) as the average effect of the tax policy is that the evolution of tax collection in municipalities that joined the program (treated municipalities) would have been the same in the absence of the program as the evolution in municipalities that never join (control municipalities) once the impact of time-varying covariates is controlled for (common trend assumption conditional on X). One can use variations in outcomes in the pre-treatment period to get a sense of whether this assumption is likely to hold. Results in Table 3 suggest that treated municipalities had been following trends similar to control ones before 1999, and that they did not experience shocks before applying to the program.

---

[18]Error correlation in the cross-section dimension of the panel could also be a concern if local governments adjust their tax policies to the actions taken by neighboring governments. Clustering at the state-year level to allow for such correlation however hardly affects the standard errors in all the specifications used below. Results are available from the author upon request.

Table 3: Determinants of program take-up

	(1)	(2)	(3)
Income	0.1252***	0.1203***	0.1435***
	(0.0406)	(0.0454)	(0.0465)
Population	0.2141	0.2852	0.2164
	(0.1973)	(0.2431)	(0.2151)
Taxes	-0.4123*	-0.5924*	-0.2544
	(0.2466)	(0.3275)	(0.2530)
Agr\ GDP	-0.0873***	-0.1087***	-0.1104***
	(0.0211)	(0.0235)	(0.0250)
Serv\ GDP	-0.0005	0.0048	-0.0090
	(0.0146)	(0.0168)	(0.0177)
Education	0.9979**	1.1272**	0.7256
	(0.4659)	(0.5188)	(0.5423)
Urban pop.	0.0058***	0.0063***	0.0063***
	(0.0017)	(0.0019)	(0.0020)
Inequality	0.0028	0.0052	0.0038
	(0.0044)	(0.0051)	(0.0050)
Distance to closest $PMAT$	-0.0032**	-0.0032*	-0.0035**
	(0.0015)	(0.0019)	(0.0017)
Time	0.0020**	0.0066***	0.0023**
	(0.0010)	(0.0011)	(0.0011)
Time^2	-0.0003***	-0.0006***	-0.0003***
	(0.0001)	(0.0001)	(0.0001)
Governor's party	-0.0003	-0.0005	-0.0004
	(0.0006)	(0.0006)	(0.0006)
Pol. competition	0.0048**	0.0050**	0.0052**
	(0.0020)	(0.0023)	(0.0023)
Lagged growth in GDP		0.0303	
		(0.0873)	
Lagged growth in population		-5.7851	
		(5.7434)	
Lagged growth in taxes		-0.0000	
		(0.0000)	
Change in mayor	-0.0012	-0.0008	
	(0.0008)	(0.0011)	
Growth in GDP 96-99			-0.7615
			(0.5008)
Growth in population 96-99			-0.3873
			(2.3233)
Growth in taxes 96-99			-0.4153
			(0.2858)
Observations	27845	23721	25040
Municipalities	3370	3349	3043

17

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Another way to detect pre-treatment trends is to take an event-study approach and examine differences between treated and control municipalities before the onset of the program. I therefore estimate a more flexible version of equation (14):

$$Y_{i,t} = \sum_{j=-9}^{9} \beta_j P_{jit} + \delta X_{i,t} + \gamma_t + \mu_i + \epsilon_{i,t}. \tag{15}$$

Here P_{jit} is equal to 1 if municipality i is in the jth year of the program in year t if $j \geq 0$, or if the municipality will sign a program contract in j years if $j < 0$. The β_j estimates are of interest for two reasons. First, when $j > 0$ they can be interpreted as estimates of the impact of the program in the jth year and measure to what extent effects are sustainable over time. Second, testing for pre-treatment trend is equivalent to a test that the β_j are equal to zero for $j < 0$. The omitted dummy is the dummy for starting a program in two years. The program's rules specify that expenditures on tax administration undertaken up to six months before the contract is signed are eligible for reimbursement so we could see a small impact of the program the year before its official start.

A final concern arises if pre-treatment characteristics potentially correlated with the dynamics of the outcome variable are unbalanced between treated and control municipalities. Convergence in tax revenues over time may, for example, lead to different dynamics between the two groups.[19] In this case difference-in-differences estimates may suffer from two additional sources of bias (Heckman *et al.* (1998)). The first occurs when there are no comparable control municipalities for some of the municipalities that join the program. The second arises from different distributions of observable covariates in the control and treated groups. Treated municipalities are different from control ones along several observable dimensions so both these types of bias are here a concern.

I therefore complement my empirical analysis by estimating a propensity score-weighted version of equation (14) following Hirano and Imbens (2001) (see also Hirano *et al.* (2003)). Propensity score-weighted regression methods eliminate both sources of bias by 1) restricting the sample to observations in the common support in the distribution of covariates, and 2) obtaining balance of covariates by re-weighting the control group observations.[20] Hirano *et al.* (2003) show that this estimator is efficient and Wooldridge (2007) that ignoring the first-stage estimation of the selection probabilities when performing inference yields conservative standard errors. All results below present standard-errors non-adjusted for first stage estimation, as bootstrapping procedures suggest there is little efficiency lost in doing so (results not shown). More details on the construction of the weights and the common support sample is found in the Appendix.

Descriptive statistics

Table 4 presents summary statistics of key characteristics of treated and control municipalities in the whole sample (column 2) and the weighted common support sample (column 3). I consider

---

[19]This could be addressed by interacting pre-treatment covariates with a time trend, but restricting their effect to be linear may not be suitable if the treatment effect is heterogenous (Meyer (1995))

[20]In practice this is done by estimating a model of the probability that a municipality joins the program as a function of the set of covariates W used in Table 3, obtaining the predicted probability $\hat{P}(W)$ and then estimating (14) with weights equal to unity for the treated and $\hat{P}(W)/(1 - \hat{P}(W))$ for the controls.

pre-program values whenever possible. Municipalities that eventually join the program are different from the average municipality when we consider the outcome variables of interest, as expected from the analysis of determinants of selection. They levy more taxes, have more municipal health infrastructure (number of health establishments per 100,000 inhabitants managed by the municipal government), less education infrastructure (number of classrooms in use in municipal schools) but of better quality. They are also found to be less corrupt over the three years for which there is data available from the audits (these cover pre and post program periods). The last column shows that restricting and weighting the sample of control municipalities leads to a reasonable balance in pre-treatment characteristics. The complete distribution of covariates in treated and control groups is also of interest. Appendix Figures 11, 12 and 13 compare the distribution of characteristics of control and treated municipalities. They show that weighting and restricting the sample to the common support makes the entire distributions of pre-treatment covariates in control municipalities very similar to that of the treated ones.

Results

Figure 3 presents a graphical representation of the results from the estimation of equation (15). Each point on the solid lines summarizes the effect of having been in the program for j years (for j positive ordinate values) or of starting the program in j years (for j negative ordinate values) compared to the year just before the program started. The excluded dummy is that for 2 years prior to the onset of the program. We see that there is no evidence of different trends prior to the onset of the program. Tax revenues start increasing as soon as the program starts, up to nearly 20 extra Rs per capita after 5 years in the program.

Table 5 reports results from the estimation of equations (14) and (15). The first two columns presents results for a model estimated on the whole sample, the third and fourth columns results obtained on the common support sample and the fifth and sixth columns results from the specification using weights based on the estimated propensity score. Estimates of equation (15) in columns 2, 4, 6 and 7 are estimated on the sample of control municipalities and treated municipalities that start a program in 2003 and 2004 only to obtain estimates of the dynamic impact of the program that cover four years before the start of the program and 5 years after.[21] Finally column 7 presents estimates of the impact of the program interacted with the per capita amount lent to estimate the returns to investment in tax capacity.

---

[21]The specifications include the full set of β_j in equation (15) so that the only excluded dummy is that for two years before the start of the program.

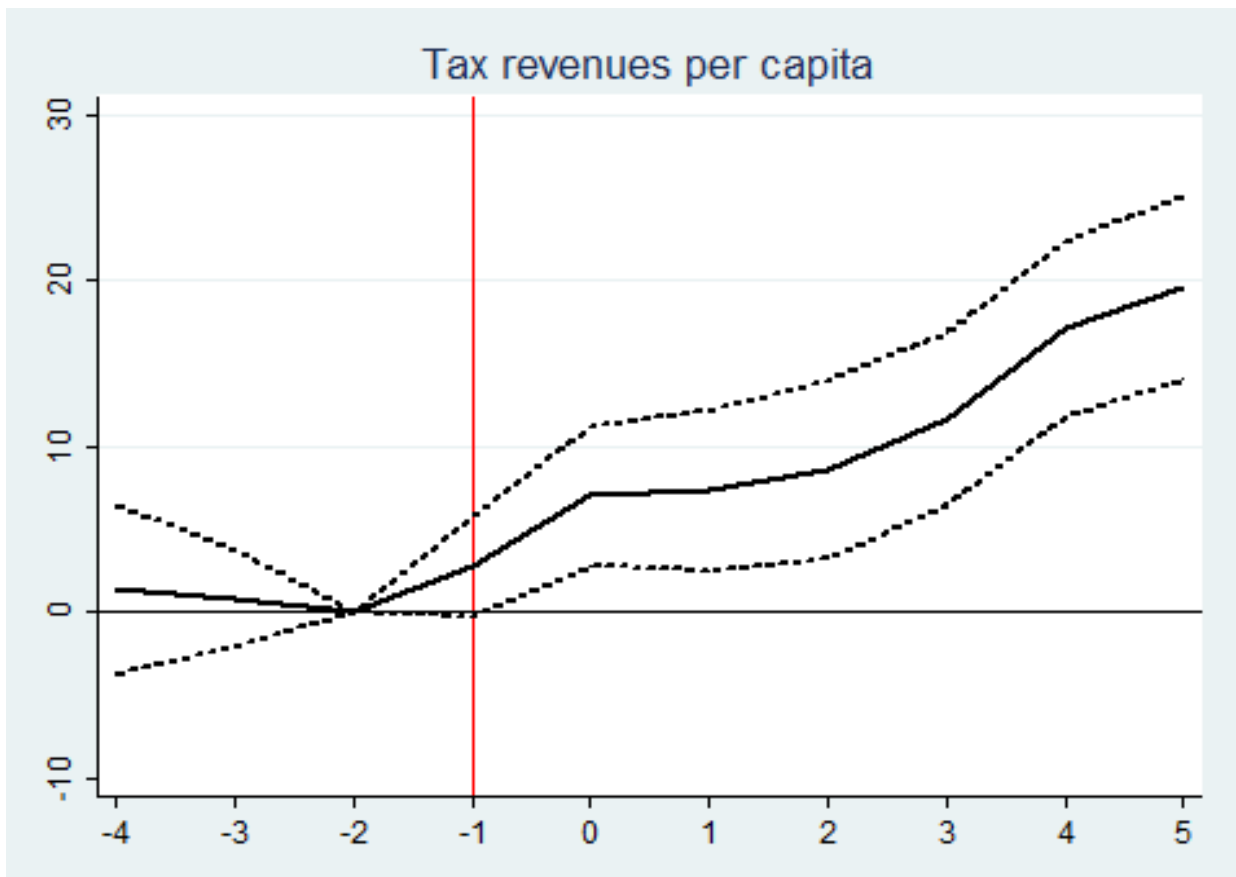Figure 3: Year by year impact of the program

Tax revenues per capita

Table 4: Descriptive statistics

	Treated	Controls	Controls, weighted
Taxes per capita (1998)	85.43	58.61	78.08
	(89.79)	(113.4)	(113.1)
FPM transfers per capita (1998)	82.03	103.1	83.87
	(43.37)	(113.4)	(51.15)
Education expenditures (% total, 1998)	30.5	32.7	29.1
	(7.66)	(7.53)	(9.79)
Health expenditures (% total, 1998)	18.9	18	20.1
	(6.63)	(6.81)	(9.54)
Quantity of education infrastructure (1998)	2.996	3.860	2.928
	(1.493)	(2.017)	(1.398)
Quality of education infrastructure (1998)	-0.217	-0.715	-0.723
	(1.311)	(1.314)	(1.204)
Health infrastructure (1999)	12.61	6.396	13.38
	(11.23)	(7.060)	(22.19)
Corruption index (all years)	57.23	184.60	46.88
	(12.2)	(432.66)	(85.21)
Public revenues per capita (1998)	394.1	415.1	394.9
	(171.0)	(205.8)	(201.2)
Population (1998)	74714.8	28457.5	82409.8
	(104565.4)	(58798.4)	(175336.4)
GDP per capita in (1999)	5972.4	4916.4	5144.9
	(4771.8)	(4114.6)	(4325.2)
Agriculture (% GDP) (1999)	11.83	22.25	21.35
	(12.54)	(15.53)	(18.70)
Services (% GDP) (1999)	62.13	59.25	61.17
	(12.24)	(13.77)	(15.57)
Income per capita (2000)	3.280	2.541	2.836
	(1.166)	(1.067)	(1.672)
Gini (2000)	0.547	0.553	0.564
	(0.0528)	(0.0563)	(0.0656)
Median education level (2000)	5.410	4.605	4.790
	(1.046)	(1.139)	(1.679)
Life expectancy (2000)	71.16	69.68	69.13
	(3.254)	(3.930)	(4.685)
Has a local radio station (1998)	0.622	0.317	0.356
	(0.486)	(0.465)	(0.479)
Has internet (1998)	0.590	0.242	0.344
	(0.493)	(0.429)	(0.475)
Has judiciary branch (1998)	0.771	0.588	0.428
	(0.421)	(0.492)	(0.495)

The first column presents averages for all treated municipalities (331, of which 279 in the common support sample), the second averages for all control municipalities (3132), the third averages for all control municipalities in the common support sample (2341), where each municipality is assigned a weight proportional to its estimated probability of joining the program.

Table 5: Impact of the program on tax revenues

	DiD		DiD on common support		Weighted DiD		Amount Paid	Treated only
All years	10.284***		9.464***		9.906***			5.353**
	(2.174)		(2.771)		(1.929)			(2.732)
4 years before		-0.952		-0.111		1.341	0.097	
		(2.966)		(3.535)		(3.098)	(0.288)	
3 years before		-1.050		0.982		0.828	0.026	
		(1.762)		(2.073)		(1.771)	(0.188)	
1 year before		-1.371		1.030		1.733	0.262	
		(1.538)		(1.668)		(1.830)	(0.191)	
1st year		3.454		4.226		3.983*	0.783***	
		(2.669)		(2.914)		(2.258)	(0.308)	
2nd year		5.229		6.521*		6.307**	0.880***	
		(3.162)		(3.383)		(2.136)	(0.304)	
3rd year		6.375*		7.023*		8.574***	1.076***	
		(3.429)		(3.638)		(3.267)	(0.359)	
4th year		10.397***		9.523**		11.644***	1.316***	
		(3.550)		(3.731)		(3.150)	(0.349)	
5th year		13.622***		14.174***		15.053***	1.754***	
		(3.681)		(3.863)		(3.213)	(0.359)	
Observations	40268	38560	28132	26818	28112	26798	26798	1614

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables is real tax revenues per capita. All regressions include municipality and year fixed effects and controls for GDP per capita, population size, share of agriculture and services in GDP, political competition in the previous election, mayor's party affiliation and whether the mayor is a facing a term limit.

The estimates suggest the program increases tax revenues by 9 to 10 Rs per capita on average, a 11% increase with respect to the baseline level. We see no different trend prior to the onset of the program. Estimates from the preferred weighted difference-in-differences specification in column 6 show a small increase in tax collection the year just before the start of the program. As explained before this may be due to actions undertaken six months before the contract was signed and eligible for reimbursement by the program's loans. It takes four years for tax revenues to increase by 10 Rs and it seems the long-run effect of the program on tax collection may be larger as it keeps increasing after four years. The estimates of the impact of 1 Rs lent through the program in column 6 suggests it is very cost-effective: after three years in the program one Rs lent leads to more than one extra Rs in tax revenues every year.

One could alternatively only use only municipalities which have not joined at time t but will later at a time $t + s$ as a counterfactual group, as those are arguably very similar to the municipalities which have already joined at time t. The bunching of municipalities' program start date around a few main years makes it impossible to estimate equation (14) on a sample consisting only of the 331 municipalities who enter the program before 2009 and identify separately year fixed effects and the program's impact. The last column of Table 5 nevertheless presents results from a similar exercise, namely the estimation of equation (14) on the sample of 331 ever-treated municipalities for the years 1999 to 2003. In this sample the municipalities which start the program after 2003 are never observed in the program, and are therefore used as control municipalities for those which enter up to 2003. Treated municipalities in this sample have been in the program on average only 1.7 years. The coefficient for the program's impact is close to the estimated impact of having been 1 or 2 years in the program using the preferred specification.

Disentangling selection and program effects

One concern remains regarding the interpretation of the estimates above as the impact of the program: I cannot completely rule out the possibility that unobservable shocks determine selection in the program and increase tax collection. The arguably exogenous time lag between program application and program start does however allow me to look for a potential selection effect: selection should affect tax collection from the date at which a municipality applies *even if the program itself does not start for a couple of years*. Appendix Table 14 presents estimates of the average impact of the program and changes during the 3 years prior to the start of the program for municipalities which apply and start a program in the same year or wait one, two or three years.It shows that tax revenues do not increase before the start of the program in any of the groups. In particular they do not increase more for municipalities which signal their willingness to invest in their tax administration a couple of years before they start the program.[22]

Finally, a recurring criticism of the difference-in-differences methodology is that it is strongly functional form dependent (Heckman, 1996). Appendix Table 13 presents results for the dynamic impact of the program using a log specification or on the whole sample of Brazilian municipalities (results for tax revenues are in the first two columns). The results are similar to what we obtain with the main specification.

---

[22]The size of the coefficients suggests a (statistically insignificant) impact of the program the year before it starts in municipalities that had already applied at that time probably because they anticipated the start of their contract.

The fact that a tax capacity program successfully increases tax collection is of interest in itself, and this last set of results suggests that increased motivation of the local government (signalled by the timing of application to the program) is not a sufficient condition for this observed success. However confidence with respect to its causal impact does not imply that the program would lead to such outcomes if applied to all Brazilian municipalities. Local motivation is likely to be a necessary condition for the program to increase tax revenues: we should interpret the 11% increase in tax collection as the program's average treatment effect on the treated (ATT), and keep this distinction in mind when comparing how municipalities spend increases in tax and transfer revenues in the next section.[23]

4.2   The transfer policy

The allocation rule described in the previous section has been used in previous papers to estimate the impact of one more unit of municipal revenues by implementing a fuzzy regression-discontinuity design strategy, comparing municipalities just below and just above thresholds (see Brollo *et al.* (2012), Litschig and Morrison (2010), Litschig (2011)). An alternative is to consider only within-municipality variations, ie to estimate the impact of transfers using municipalities which pass from one population bracket to the next over time. The strategy used to identify the impact of one more unit of tax revenues used in this paper relies on using only within-municipality variations (or within group variations over time when the outcome variable is corruption) and I wish to compare the impact of taxes and transfers using the same specification. There is no a priori reason for within municipality and between municipalities estimates to differ. In this section I show that the impact of the rule on transfer revenues is, as expected, very similar when we consider variations within and between municipalities.

This identification strategy is valid if 1) there are no other policies that are discontinuous at the same population thresholds and 2) population I use to construct theoretical transfers are not manipulated by local governments to sort above the threshold. One other policy is discontinuous in municipal population size: the wage of local councillors is capped and increases discontinuously when the population reaches 10,000 50,000 and 100,000 inhabitants (see Ferraz and Finan (2008b)). Three *FPM* thresholds are close to those (thresholds at 10,188 50,940 and 101,880 inhabitants), it turns out that results are very similar if we exclude these thresholds (results available from the author upon request).

The population variable I use as running variable comes from annual estimates by the Brazilian statistical agency (*IBGE*). The Appendix describes how these estimates are obtained and why it is very unlikely that municipalities could manipulate them.[24] Appendix Figure 6 checks visually for evidence of manipulation of the running variable by showing the frequency of municipalities

---

[23]The weights used in the weighted difference-in-differences specification are appropriate to compute the average treatment on the treated effect only (Hirano and Imbens, 2001). It seems fallacious to try to estimate an average treatment effect for the whole population of Brazilian municipalities given the great heterogeneity of fiscal and economic contexts of Brazilian municipalities and the fact that cooperation of local tax administration is by definition necessary for the program to work.

[24]The *Tribunal de Contas União* is supposed to use these estimates to determine how much each municipality will receive. However it publishes the population estimates it effectively uses in its computations every year and they do not perfectly coincide with the *IBGE* estimates used here. It is likely that any manipulation of population estimates happens after the *IBGE* publishes its estimates and before the *TCU* announces tranfer amounts.

within population bins and running McCrary tests for the presence of a density discontinuity at the thresholds (see McCrary (2008)). I perform the tests on all thresholds simultaneously by normalizing population size as the distance from the closest threshold (with symmetric intervals around each threshold so that no municipality belongs to more than one interval), and on each threshold separately. There is no sign that municipalities manipulate the population estimates: the log-difference between the frequency to the right and to the left of all thresholds is never statistically significant, despite some (visual) suspicion of a little sorting around thresholds 4 and 14.[25] The use of within municipalities variations suggests another test of manipulation of the running variable: Appendix Figure 9 plots average population growth between years $t+1$ and t as a function of population size in year t. There is no evidence that municipalities manipulate their population estimates in the following year when this could increase their future transfers, ie when they are close to a threshold.

Figure 4: The population discontinuity and *FPM* transfers

The graph plots average *FPM* transfers per capita as a function of normalized population size (population size minus the nearest threshold value). Scatter points are averaged over 75 inhabitants intervals. The solid lines are the 95% confidence interval of the means.

---

[25]The point estimate (standard errors) for the test on all thresholds simultaneously is 0.0509 (0.0607).

Figure 4 plots average *FPM* transfers per capita received as a function of (normalized) population size. We see a clear jump of 10-20 Rs at the threshold, roughly 15% of the pre-treatment amount. To formally check the impact of the transfer policies on transfer revenues I estimate the following regression:

$$F_{i,t} = g(P_{i,t}) + \beta FPM_{i,t} + \gamma_t + \mu_i + \epsilon_{i,t}, \tag{16}$$

where $F_{i,t}$ is *FPM* transfers per capita, $g(\dot{})$ is a second order spline polynomial in normalized population $P_{i,t}$ and $FPM_{i,t}$ is equal to 1 if local population is above the nearest threshold and γ_t and μ_i are year and municipality fixed effects. Table 6 presents results from estimating (16) on the pooled sample and for each individual threshold. The first column presents results obtained by using both within and between municipalities variations (replacing municipality fixed effects by state fixed effects), the second and third using only variations within municipalities. Anticipating on the analysis in the next section the third column restricts the sample to the common support sample and weights municipalities by a function of their estimated probability of joining the PMAT program to consider the impact of the transfer policy on municipalities that are similar to those affected by the tax policy.

Transfers per capita increase by 14 Rs on average in the pooled sample. Results are extremely similar with and without municipality fixed effects. As in Figure 2 the jump is decreasing in population size, from more than 20 Rs at the first threshold, to 3 Rs at the last threshold. Note that at threshold 5 transfers per capita increase by 9.5 Rs, an amount similar to the increase in taxes per capita thanks to the *PMAT* program. Finally the estimate in column three is smaller, because it gives more weight to non-*PMAT* municipalities that are similar to the ones who joined the program which are larger than the average municipality. The average effect of the discontinuity on this re-weighted sample is close to 9 Rs, similar to the impact of the program on tax revenues.

Table 6: Impact of the population discontinuity on FPM transfers

	No fixed effects		Fixed effects		Fixed effects + weights	
	All thresholds					
Discontinuity	14.375***	(0.620)	14.495***	(0.624)	8.751***	(2.041)
Observations	27459		27459		20556	
	Threshold 1					
Discontinuity	23.030***	(1.851)	23.258***	(1.889)		
Observations	5737		5737			
	Threshold 2					
Discontinuity	19.248***	(1.527)	19.531***	(1.564)		
Observations	4484		4484			
	Threshold 3					
Discontinuity	15.569***	(1.434)	15.770***	(1.466)		
Observations	4735		4735			
	Threshold 4					
Discontinuity	13.924***	(1.059)	14.294***	(1.063)		
Observations	3744		3744			
	Threshold 5					
Discontinuity	9.475***	(1.118)	9.548***	(1.140)		
Observations	2408		2408			
	Threshold 6					
Discontinuity	6.637***	(1.353)	6.258***	(1.356)		
Observations	1372		1372			
	Threshold 7					
Discontinuity	5.728***	(1.141)	6.186***	(1.127)		
Observations	1008		1008			
	Threshold 8					
Discontinuity	7.711***	(1.206)	7.899***	(1.171)		
Observations	824		824			
	Threshold 9					
Discontinuity	3.266***	(1.103)	3.081***	(1.174)		
Observations	824		824			
	Threshold 10					
Discontinuity	5.079***	(0.932)	5.338***	(0.925)		
Observations	615		615			
	Threshold 11					
Discontinuity	4.752***	(1.543)	4.740***	(1.385)		
Observations	478		478			
	Threshold 12					
Discontinuity	4.037***	(0.713)	3.875***	(0.627)		
Observations	373		373			
	Threshold 13					
Discontinuity	2.727**	(1.095)	2.785**	(1.156)		
Observations	338		338			
	Threshold 14					
Discontinuity	5.313***	(1.882)	4.338***	(1.270)		
Observations	255		255			
	Threshold 15					
Discontinuity	1.432	(1.202)	3.337***	(0.890)		
Observations	264		264			

The specification in the first column includes state and year fixed effects, the specification in the second and third columns municipality and year fixed effects. All specifications include a second order spline polynomial in population size. Observations in the third column are weighted by a function of their estimated probability of joining the program (weights are equal to 1 for municipalities that join the program). The interval around a threshold is plus or minus 1698 inhabitants for thresholds 1 to 3, plus or minus 3396 for thresholds 4 to 8, plus or minus 5094 for thresholds 9 to 13 and plus or minus 6792 for thresholds 14 and 15. Each interval is constructed so that each municipality with population size between 8490 and 143'633 is in one and one only interval around a threshold.

5 Are tax revenues better spent than transfer revenues?

This section compares the impact on municipal public expenditure outcomes of an increase in tax revenues thanks to the tax policy to the impact of an increase in *FPM* transfers. It directly tests propositions 1 to 3 of the model which predict that an increase in taxes will lead to a higher increase in public good provision and a smaller increase in corruption than an increase in transfers. I start by explaining the strategy used to identify the marginal propensity to spend out of tax and transfer revenues on local public expenditure outcomes and then present results.

5.1 Empirical strategy and validity checks

I evaluate the impact of an increase in taxes thanks to the program or an increase in transfer revenues on public spending outcomes by estimating the following equations :

$$G_{i,t} = \pi_R R_{i,t} + \eta X_{i,t} + \gamma_t + \mu_i + \epsilon_{i,t}, R = F, T \tag{17}$$

and

$$s_{i,t} = \phi_R R_{i,t} + \delta X_{i,t} + \delta_2 Z_i + \delta_3 S_i + \gamma_t + \epsilon_{i,t}, R = F, T \tag{18}$$

where $R_{i,t}$ is 1) tax revenues per capita ($T_{i,t}$) or 2) transfer revenues per capita ($F_{i,t}$), $G_{i,t}$ are measures of municipal health and education infrastructure, $s_{i,t}$ the corruption index and $X_{i,t}$ a set of time-varying covariates are as above. The corruption data is a repeated cross-section and not a panel, so specification (18) controls for an indicator of whether the municipality joins the program in the period 1998-2009 S_i and a set of time invariant covariates Z_i that includes state fixed effects and municipal characteristics.[26]

The model predicts that $\pi_T > \pi_F$ and $\phi_T < 0 < \phi_F$.
Identification of causal impacts

Different sets of assumptions are necessary to interpret the comparison of π_T and π_F and ϕ_T and ϕ_F in equations (17) and (18) as a test of the model's predictions. I instrument tax revenues by participation in the *PMAT* program and transfer revenues by the theoretical transfers constructed by applying the allocation rule to population estimates. The previous section has shown that these instruments have an economically meaningful and statistically significant impact on tax and transfer revenues. The exclusion restriction for the instrument for transfers requires that no other policy is discontinuous at the thresholds and that no municipality is able to manipulate the population estimate used to construct theoretical transfers. Both of these assumptions have been discussed in the previous section.

The methods used above to consider whether the program has a causal impact of tax revenues are also relevant to assess the plausibility of the exclusion restriction for the instrument for tax revenues, namely that participation in the program has no impact on expenditure outcomes beyond increasing tax revenues. Figure 5 shows coefficients obtained from estimating equation (15) using municipal education quantity and quality as outcomes. We see that there is no different

---

[26]These are median education level, inequality and life expectancy from the 2000 Census, whether tourism is a major industry, existence of a local radio station, local judiciary presence from the 1998 *Perfil dos Municipios Brasileiros* and whether the municipality is a state capital.

trend in program participants prior to the start of the program and a gradual increase once the program starts. Appendix table 15 shows more generally that expenditure outcomes are not following different trends prior to the start of the program.[27] Appendix Table 16 similarly shows that having self-selected into the program but not started one yet has no impact on the quantity and quality of municipal education infrastructure. This suggests that any impact of higher tax revenues thanks to the program on expenditure outcomes is unlikely to be due to a selection effect.

The exclusion restriction also requires that the program must not have a direct impact on public spending outcomes beyond increasing tax revenues. The program's funds cannot be spent on anything but modernizing the tax administration. Even assuming municipalities manage to bend this rule and spend the funds on education and health infrastructure a simple back of the envelope calculation suggests it is highly unlikely that the program's small loans was sufficient to fund the increase in infrastructure observed in Figure 5. Treated municipalities had roughly 400 Rs per capita of total public revenues in 1998 and 3 classrooms in use per 1000 inhabitants. Using the average propensity to spend on classrooms these municipalities would need an extra 70 Rs to open 0.2 new classrooms, the estimated effect of the program. The average loan amount, 9 Rs per capita, seems way too small to finance such an increase. The cumulated increase in tax revenues after 4 years in the program, nearly 40 Rs, plausibly can if we assume that the marginal propensity to spend on classrooms is slightly higher than the average propensity to spend.

---

[27]I restrict all pre-treatment dummies to be equal when health infrastructure is used as an outcome because I have observations for few pre-treatment years for this variable.

Figure 5: Year by year impact of the program
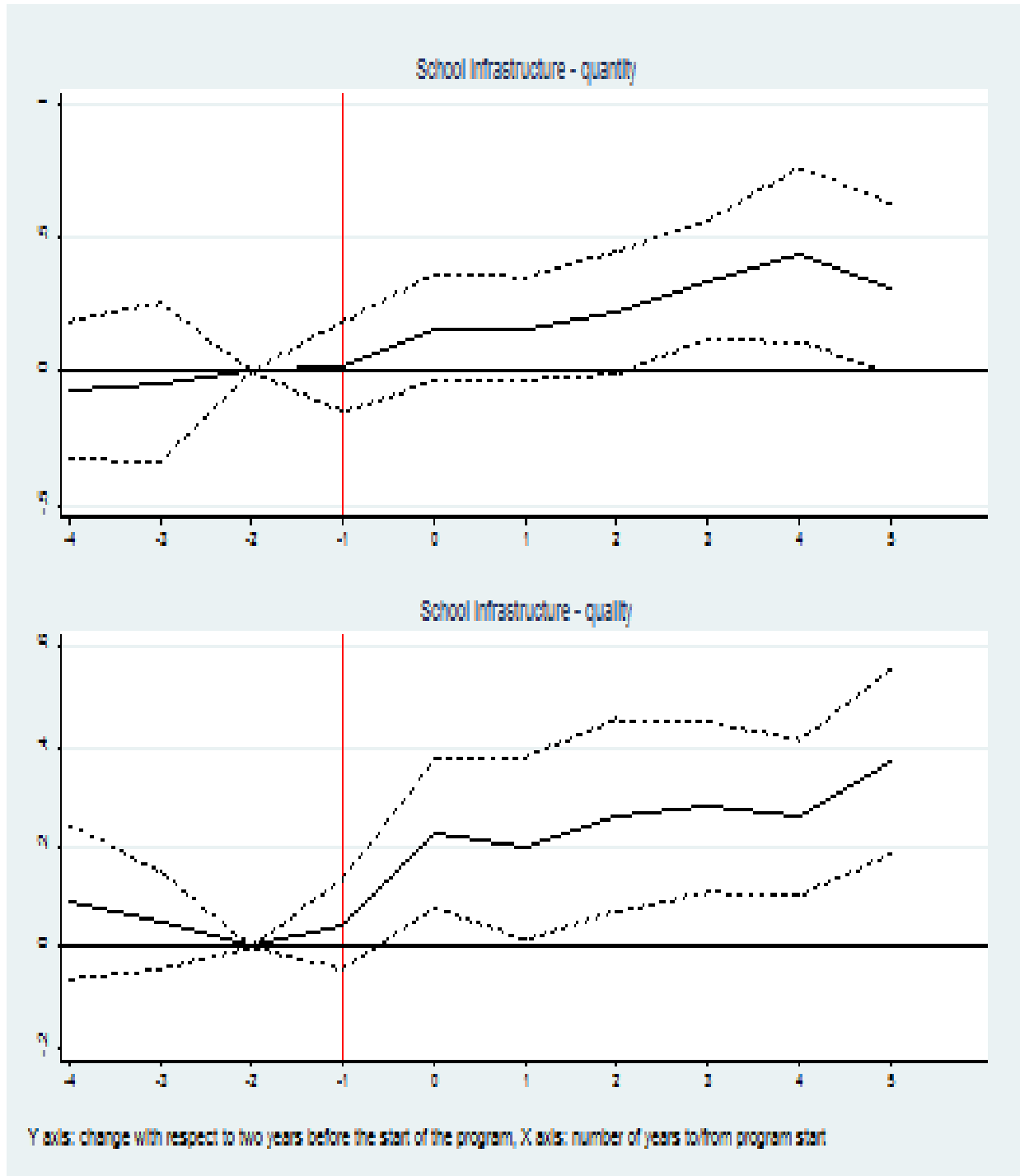
The graphs plot the coefficients from estimating equation (15) on a sample of control municipalities and municipalities treated in 2003, 2004 and 2005. The dependent variable is the number of classrooms in use in municipal schools per 1000 inhabitants in the top panel and the index of municipal school quality in the bottom panel.

30

5.2 Results

Table 7 presents estimates of the impact of a 10 Rs increase in taxes and transfers on municipal public spending outcomes. The first column presents results for the impact of transfers in a specification without municipality fixed effects but with the full set of time-invariant controls. Estimates in the second and third columns are obtained from running specification (17) (top three panels) and (18) (bottom panel) in the full sample.[28] The last two columns present estimates obtained from running these specifications on the weighted sample.

An increase in transfers leads to a small improvement in both the quantity and quality of municipal education infrastructure in the unweighted sample, but these effects are much smaller than the impact of an increase in taxes. They disappear in the specification using weights and restricting the sample to the common support sample, suggesting the impact of transfers is smaller in municipalities that are similar to those that join the program. An increase of 10 Rs in tax revenues thanks to the program leads to an extra 0.18 classrooms per thousand inhabitants, a nearly 6% increase compared to an average 3.1 classrooms in treated municipalities the year before the program starts. It increases the school quality index by 0.2, a small amount corresponding roughly to one-fifth of a standard deviation. The impact on the number of health establishments per 100,000 inhabitants is more substantial, it increases by nearly 7% compared to its pre-program level. Higher transfers have no impact on health infrastructure. Finally, higher transfers lead to a large increase in the corruption index of 10 to 12 (6% compared to the unweighted average, 23% compared to the weighted average). This is in line with the results in Brollo *et al.* (2012) that a 10% increase in transfers leads to a 17% increase in corruption. The increases in taxes has a negative but not statistically significant impact on the corruption index, which suggests the index may have decreased by 10 to 15% compared to its pre-program level in the 20 municipalities audited before they started a program. The results are therefore in line with the predictions of the model.

Interpreting results in Table 15 as a validation of the structural difference between predicted by the model between π_T and π_F and ϕ_T and ϕ_F requires making an additional assumption regarding the heterogeneity of the effects. To see this note that the estimates are local average treatment effects (LATE): they represent the average marginal propensity to spend out of taxes (transfers) amongst municipalities that are affected by the tax (transfer) policy. Estimated propensities to spend out of taxes and transfers could therefore differ if municipalities spend taxes and transfers in the same way ($\pi_F = \pi_T = \pi$, $\phi_F = \phi_T = \phi$) but there is heterogeneity in their marginal propensities to spend out of increases in all types of revenues, and π and ϕ are different in municipalities affected by the two instruments. For example we might think that the marginal propensity to spend public revenues on education infrastructure is higher in municipalities that have less public revenues to start with and less education infrastructure.

The fact that we can estimate 15 different local effects of an increase in transfers on different sub-samples around the 15 thresholds helps assess the extent to which the marginal propensity to spend out of transfers does varies in the population. Appendix Table 11 shows the impact of extra transfers on health and education infrastructure for each threshold separately.[29] We see

---

[28]When transfers are on the right-hand-side the sample includes only municipalities with more than 8400 and less than 142'633 inhabitants and the specification includes a spline polynomial.

[29]I cannot repeat the exercise using the corruption index as the sample size is too small.

Table 7: Marginal propensity to spend out of a 10 Rs increase in taxes or transfer revenues (IV estimates)

	No fixed effects	Fixed effects		Fixed effects + weights	
Dependent variable: quantity of municipal education infrastructure					
Transfers per capita	0.023**	0.024		0.003	
	(0.012)	(0.033)		(0.030)	
Taxes per capita			0.161***		0.180***
			(0.050)		(0.053)
Observations	27438	27393	40373	20544	28214
Dependent variable: quality of municipal education infrastructure					
Transfers per capita	0.034***	0.055***		0.007	
	(0.008)	(0.008)		(0.023)	
Taxes per capita			0.228***		0.222**
			(0.071)		(0.113)
Observations	27420	27375	40335	20526	28182
Dependent variable: quantity of municipal health infrastructure					
Transfers per capita	-0.137	-0.100		-0.168	
	(0.111)	(0.168)		(0.179)	
Taxes per capita			1.085**		1.219*
			(0.519)		(0.641)
Observations	9697	9579	14255	7226	9999
Dependent variable: corruption index					
Transfers per capita	10.751***	12.832***			
	(3.693)	(4.490)			
Taxes per capita		-10.158			
		(9.607)			
Observations	424	509			

Cluster-robust standard errors in parentheses. All specifications include year fixed effects and time-varying controls (GDP per capita, population size, the shares of agriculture and services in GDP, whether the mayor is in his second term, and political competition in the last election). Specifications without fixed effects (first column and last panel) include state fixed effects and as controls a indicator of whether the municipality joins a *PMAT* program between 1998 and 2008 and municipal characteristics (income per capita, inequality, median education, life expectancy, population density, all in 2000, and whether the municipality has a local radio station in 1998 and a seat of the judiciary branch in 1998). When transfers per capita are the main explanatory variable the sample includes only municipalities with more than 8400 and less than 143000 inhabitants and the specification includes a second order spline polynomial.

that, whilst coefficients vary at each threshold they always remain lower than the average impact of an increase in taxes. The increase in transfers has as expected a larger impact on the quantity of school infrastructure in smaller municipalities: 0.07 to 0.1 extra classrooms at thresholds 1 to 4. This is coherent with the results in Litschig (2008) that an increase in transfers leads to better education outcomes amongst small municipalities. The largest point estimate (not statistically significant) is found at threshold 7: at 0.145 it is still 20% smaller than the estimated impact of tax revenues.[30] Assuming that these estimates are plausible bounds for the underlying heterogeneity of the parameters of interest allows us to interpret the difference between the estimates as a sign that municipalities do spend tax and transfer revenues differently.

The re-weighting of the specification in the last two columns of Table 7 ensures that the aggregation of these 15 estimates gives more weight to the sub-samples that are more similar to municipalities affected by the tax policy. Appendix Table 10 presents descriptive statistics for each sub-sample affected by a different threshold as well as their average estimated weight. Thresholds 5 to 11 are given a higher weight in the estimation, threshold 5 is of particular interest as it experiences an increase in transfers that is of roughly the same size as the increase in taxes. Municipalities affected by thresholds 8,9 and 10 look particularly similar to municipalities that join the program in terms of revenue per capita and municipal infrastructure. The fact that the estimated propensity to spend out of transfers around these thresholds is always lower than the estimated impact of taxes provides further reassurance that different baseline levels of revenues or demand for public services are not driving the results.

5.3   Discussion

Why are tax revenues spent more on provision of public services and less on corruption than transfer revenues? The model above suggests that increases in tax revenues are better observed than increase in transfer revenues by citizens, and hence better spent. Table 8 offers a test of this mechanism. Following Ferraz and Finan (2011) I use the presence of a local radio station as a proxy for how much information citizens can access about local public budgets. If the presence of a local radio station decreases information asymmetries between citizens and elected officials (lower u in the model) we expect transfers to be spent better when there is a local radio station, and no change in the impact of taxes. This is a test of Proposition 3 above.

The coefficients for the interaction term between the radio variable and transfers per capita are in line with the model's predictions: transfers are better spent when there is a local media. Results in the last column are however somewhat surprising: most of the negative impact of tax revenues on the corruption index comes from municipalities which have a radio station (60% of municipalities in the $PMAT$ program). This suggests that transfer revenues are not the only parts of the public budget imperfectly observed by citizens - if the cost of producing public services is also badly observed by citizens for example we expect better information to improve how both tax and transfer revenues are spent.

---

[30]The estimated impact on the quality of education infrastructure similarly varies between -0.16 and 0.16 and is statistically significant in some smaller municipalities (thresholds 1 2 and 3) as well as some bigger ones (thresholds 7,9 and 12). There seems to be no impact of higher transfers on municipal health infrastructures at any of the population thresholds - coefficients tend to be negative.

Table 8: Marginal propensity to spend out of taxes and transfers with and without a local radio station

	Educ. infrastructure: quantity		Educ. infrastructure: quality		Health infrastructure		Corruption	
Transfers per capita	-0.007		0.010		-0.179		10.851***	
	(0.030)		(0.013)		(0.191)		(3.657)	
Transfers*Radio	0.009		0.028*		0.112*		-1.015	
	(0.006)		(0.015)		(0.059)		(2.015)	
Taxes per capita		0.217		0.232*		1.101***		1.761
		(0.135)		(0.117)		(0.314)		(11.210)
Taxes*Radio		-0.011		-0.003		0.046		-11.108
		(0.092)		(0.021)		(0.103)		(20.162)
Radio							-8.581	-7.627
							(21.754)	(12.684)
Observations	20544	28214	20526	28182	7226	9999	424	526

Cluster-robust standard errors in parentheses. All specifications include year fixed effects and time-varying controls (GDP per capita, population size, the shares of agriculture and services in GDP, whether the mayor is in his second term, and political competition in the last election). Specifications without fixed effects (last two columns) include state fixed effects and as controls a indicator of whether the municipality joins a *PMAT* program between 1998 and 2008 and municipal characteristics (income per capita, inequality, median education, life expectancy, population density, all in 2000, and whether the municipality has a local radio station in 1998 and a seat of the judiciary branch in 1998). When transfers per capita are the main explanatory variable the sample includes only municipalities with more than 8400 and less than 143000 inhabitants and the specification includes a second order spline polynomial.

Several other mechanisms could lead to a difference in how increases in tax and transfer revenues are spent. First, it could be that governments which rely more on local tax revenues have better incentives to invest in public goods if this increases the local tax base, in a reformulation of Tiebout (1956). It is unlikely that the types of health and education investments local governments can make in Brazil affect local growth fast enough to be a relevant factor for politicians (and detected in my sample) but better public services may attract more taxpayers even in the short-run. I find no evidence that the tax policy has any impact on local GDP or population (Appendix Table 17) but cannot rule out the possibility that this mechanism plays a role. Second, it may be that local governments only invest in infrastructure when they experience an increase in revenues that they believe is stable over time, and divert increases in revenues that are short lived. This could explain the results above if tax revenues generated by the tax policy are more stable than transfer revenues. The within-municipality standard deviation is however always smaller relative to the mean for transfer (FPM) revenues than for tax revenues (Appendix Table 18): tax revenues are not less volatile than the type of transfers considered so different volatility levels cannot explain the results.[31] Finally, the literature in behavioral economics argue that individuals care more about out-of-pocket costs than about opportunity costs (see Hines and Thaler (1995) for an application of this idea to the fly-paper effect). This implies that citizens should care more about how tax revenues (out-of-pocket costs) are spent than about transfer revenues (from their point of view, opportunity costs) and hence that they may be more willing to monitor politicians when they pay more taxes. This last explanation cannot be investigated in the absence of a good measure of citizen monitoring activities (voting is compulsory in Brazil and there is little variation in turnout across municipalities).

---

[31]This is unsurprising - FPM transfers only vary if the total amount allocated to FPM transfers at the federal level changes or if the municipality's population reaches a threshold. Discretionary transfers or aid revenues are likely to be much more volatile.

6 Conclusion

This paper shows that a local tax modernization program in Brazil leads to a permanent 11% increase in taxes per capita, an increase in municipal health and education infrastructure, and no increase in the incidence of a broad measure of corruption. I take advantage of the variation in taxes induced by the program and discontinuities in the rule allocating federal transfers to test a theoretical prediction that taxes are more *accountability inducing* than transfers. Results show that local governments use the increase in taxes thanks to the program to provide more health and education infrastructure than they do when faced with an increase in transfer revenues of the same amount. More transfers lead to more corruption, more taxes do not.

These results speak directly to debates about the right form of decentralization. The existence of a large 'fiscal gap' between local expenditure responsibilities and local tax revenues is an ubiquitous characteristic of local governments around the world. In developing countries in particular these governments have been granted substantial expenditure responsibilities, but local capacity to tax generally lags behind. My results suggest that narrowing this fiscal gap by empowering local governments to levy more tax revenue will make them more accountable to their constituents. Substantial local tax collection – complemented by intergovernmental transfers for revenue equalization purposes – is a necessary feature of successful decentralization.

Moving up from the local government level the mechanisms explored in this paper also contribute to debates on how to finance development. One of the recommendation of the 2005 report on achievement of the Millenium Development Goals is that developing countries should increased domestic resources by up to four percentage points by 2015 (Sachs *et al.*, 2005). There is however very little research on how this aim could be achieved. Moreover, technical aid on public sector financial management has always been the poor parent of official development aid.[32] This paper shows that one type of resource mobilization program in place in Brazil for more than a decade has been successful in providing long term sources of funds to local governments. Moreover, the theoretical argument developed in this paper also applies to a federal government financed by tax and non-tax revenues (such as aid or revenues from natural resources). It suggests that technical help in tax capacity building may lead to an increase in public resources which is more conducive to the type of public spending that benefits citizens than traditional development aid.

---

[32]See OECD (2010a) for a discussion of the different forms of aid in public sector financial management.

References

AFONSO, J. R. (2005). Decentralization and budget management of local government in brazil, world Bank.

— and ARAUJO, E. (2006). Local government organization and finance : Brazil. In A. Shah (ed.), *Local Governance in Developing Countries*, World Bank.

— and SERRA, J. (1999). Fiscal federalism brazilian style: Reflections, mimeo.

BARDHAN, P. and MOOKHERJEE, D. (2006). *Decentralization and Local Governance in Developing Countries*. MIT Press.

BESLEY, T. (2006). *Principled Agents? The Political Economy of Good Government*. Oxford University Press.

— and PERSSON, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, **99** (4), 1218–1244.

— and — (2010). State capacity, conflict, and development. *Econometrica*, **78** (1), 1–34.

— and — (2011). *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press.

— and SMART, M. (2007). Fiscal restraints and voter welfare. *Journal of Public Economics*, **91** (3-4), 755–773.

BNDES (2002). *Modernizacao da Gestao Publica: Uma Avaliacao de Experiencias Inovadoras*. Banco National do Desenvolvimento.

BRAUTIGAM, D., FJELDSTAD, O.-H. and MOORE, M. (eds.) (2008). *Taxation and State-Building in Developing Countries*. Cambridge University Press.

BROLLO, F., NANNICINI, T., PEROTTI, R. and TABELLINI, G. (2012). The political resource curse. *American Economic Review*, **Forthcoming** (15705).

CAGÉ, J. and GADENNE, L. (2012). *The Fiscal Cost of Trade Liberalization*. PSE Working Papers halshs-00705354, HAL.

CASELLI, F. and MICHAELS, G. (2011). *Do Oil Windfalls Improve Living Standards? Evidence from Brazil*. Working Papers 15550, National Bureau of Economic Research.

DE CARVALHO JR., P. H. B. (2006). *IPTU no Brasil: Progressividade, Arrecadao e Aspectos Extra-Fiscais*. Discussion Papers 1251, Instituto de Pesquisa Econmica Aplicada - IPEA.

DE PAULA, u. and SCHEINKMAN, J. A. (2010). Value-added taxes, chain effects, and informality. *American Economic Journal: Macroeconomics*, **2** (4), 195–221.

EASTERLY, W. (2008). *Reinventing Foreign Aid*. Cambridge MA: MIT Press.

FERRAZ, C. and FINAN, F. (2008a). Exposing corrupt politicians: The effects of Brazil's publicly released audits on electoral outcomes. *The Quarterly Journal of Economics*, **123** (2), 703–745.

— and — (2008b). *Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance*. IZA Discussion Papers 3411, Institute for the Study of Labor (IZA).

— and — (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, **101** (4), 1274–1311.

—, — and MOREIRA, D. B. (2012). *Corrupting Learning:Evidence from Missing Federal Education Funds in Brazil*. Working Paper 18150, National Bureau of Economic Research.

— and MONTEIRO, J. (2010). Does oil make leaders unaccountable? evidence from brazil's offshore oil boom, paper presented at the NEUDC Conference, November 2010.

FISMAN, R. and GATTI, R. (2002). Decentralization and corruption: Evidence from u.s. federal transfer programs. *Public Choice*, **113** (1-2), 25–35.

GALIANI, S., GERTLER, P. and SCHARGRODSKY, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *Journal of Political Economy*, **113** (1), 83–120.

GERVASONI, C. (2010). A rentier theory of subnational regimes: Fiscal federalism, democracy, and authoritarianism in the argentine provinces. *World Politics*, **62 (2)**, 302–340.

HECKMAN, J. J. (1996). Labor supply and the economic recovery tax act of 1981: Comment. In M. Feldstein and J. M. Poterba (eds.), *Empirical Foundations of Household Taxation*, University of Chicago Press, pp. 32–38.

—, ICHIMURA, H. and TODD, P. (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies*, **65** (2), 261–94.

HINES, J. and THALER, R. H. (1995). Anomalies: The flypaper effect. *The Journal of Economic Perspectives*, **9 (4)**, 217–226.

HIRANO, K. and IMBENS, G. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes Research Methodology*, **2**, 259–278.

—, IMBENS, G. W. and RIDDER, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, **71** (4), 1161–1189.

JENKINS, S. P. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics*, **57** (1), 129–38.

LITSCHIG, S. (2008). *Rules vs. Political Discretion: Evidence from Constitutionally Guaranteed Transfers to Local Governments in Brazil*. Working Papers 1144, Department of Economics and Business, Universitat Pompeu Fabra.

— (2011). *Financing Local Development: Quasi-Experimental Evidence from Municipalities in Brazil, 1980-1991.* Working Paper 1142, Department of Economics and Business, Universitat Pompeu Fabra.

— and MORRISON, K. (2010). *Government spending and re-election: Quasi-experimental evidence from Brazilian municipalities.* Working Papers 1233, Department of Economics and Business, Universitat Pompeu Fabra.

— and ZAMBONI, Y. (2008). *The Short Arm of the Law: Judicial Institutions and Local Governance in Brazil.* Working Papers 1143, Department of Economics and Business, Universitat Pompeu Fabra.

McCRARY, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, **142 (2)** (334).

MEYER, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, **13** (2), 151–61.

MONTEIRO, J. C. and ASSUNO, J. J. (2011). Coming out of the shadows? estimating the impact of bureaucracy simplification and tax cut on formality in brazilian microenterprises. *Journal of Development Economics*, (0), –.

MOORE, M. (2007). How does taxation affect the quality of governance? *Tax Notes International*, **47**, 79.

NORTH, D. C. and WEINGAST, B. (1989). Constitutions and Commitment: Evolution of Institutions Governing Public Choice. *Journal of Economic History*.

OECD (2010a). *Do No Harm : International Support for Statebuilding.* Tech. rep., Organization for Economic Cooperation and Development.

— (2010b). Public resource mobilization and aid in africa. In *African Economic Outlook*, African Development Bank and OECD.

OLKEN, B. A. (2007). Monitoring corruption: Evidence from a field experiment in indonesia. *Journal of Political Economy*, **115**, 200–249.

PALER, L. (2012). Keeping the public purse: An experiment in windfalls, taxes, and the incentives to restrain government, mimeo, Columbia University.

POMERANZ, D. (2010). *No Taxation without Information.* Working paper, Harvard Economics Department.

REINIKKA, R. and SVENSSON, J. (2005). Fighting corruption to improve schooling: Evidence from a newspaper campaign in uganda. *Journal of the European Economic Association*, **3** (2-3), 259–267.

ROCHA, R. and SOARES, R. (2010). Evaluating the impact of community-based health interventions: Evidence from brazil's family health program. *Health Economics*, **19**, 126–158.

Ross, M. L. (2004). Does taxation lead to representation? *British Journal of Political Science*, **34**, 229–249.

Sachs, J. *et al.* (2005). Un millennium development project - investing in development : A practical plan to achievet goals., united Nations, New York.

Santos, G. A. G. D., Tavares, F. H., Reiff, L. and Correa, L. (2008). Potencial de arrecadaco municipal e o pmat. *Revista do BNDES*, **29**, 399–434.

Schumpeter, J. A. (1918). The crisis of the tax state. In R. Swedberg (ed.), *Joseph A. Schumpeter. The Economics and Sociology of Capitalism*, Princeton University Press.

Strumpf, K. S. (1998). A predictive index for the flypaper effect. *Journal of Public Economics*, **69** (3), 389–412.

Svensson, J. (2000). Foreign aid and rent-seeking. *Journal of International Economics*, **51** (2), 437–461.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, **64**, 416.

Weingast, B. R. (2009). Second generation fiscal federalism: The implications of fiscal incentives. *Journal of Urban Economics*, **65**, 279–293.

Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics*, **141** (2), 1281–1301.

World Bank (2004). *World Development Report 2004: Making Services Work for Poor People.* New York, NY: Oxford University Press.

Zhuravskaya, E. V. (2000). Incentives to provide local public goods: Fiscal federalism, russian style. *Journal of Public Economics*, **76** (3), 337–368.

A Theoretical appendix

A.1 Proof that the participation constraint is binding in the L state

The representative citizen's maximizes :

$$W(G_L, T_L, T_H, \sigma_H, \sigma_L) = (1-q)(G_L - \phi C(T_L)) + q(G_L + T_H - T_L + Z(\sigma_H - \sigma_L) - \phi C(T_H)) \quad (19)$$

subject to the following constraints, where λ_i is the lagrange multiplier associated with constraint i

$$\begin{cases} 1: & G_L \leq (\bar{F}(1-u) + T_L)(1-\alpha) + \sigma_L Z \ (\lambda_1) \\ 2: & G_L \leq \bar{F}(1+u)(1-\alpha) - \alpha T_H + T_L + \sigma_L Z \ (\lambda_2) \\ 3: & T_H \geq 0 \ (\lambda_3) \\ 4: & T_L \geq 0 \ (\lambda_4) \\ & \sigma_H \in [0,1] \\ & \sigma_L \in [0,1] \end{cases}$$

where constraint 1 is the participation constraint in the L state I am using the fact that $G_H = G_L + T_H - T_L + Z(\sigma_H - \sigma_L)$ to rewrite the participation constraint in state H (constraint 2). The first order conditions for maximization are:

$$\frac{\partial W}{\partial \sigma_H} = Zq \geq 0 \quad (20)$$

$$\frac{\partial W}{\partial \sigma_L} = Z(\lambda_1 + \lambda_2 - q) \geq 0 \quad (21)$$

$$\frac{\partial W}{\partial G_L} = 0 \Leftrightarrow \lambda_2 + \lambda_1 = 1 \quad (22)$$

$$\frac{\partial W}{\partial T_L} = 0 \Leftrightarrow \lambda_4 = (1-q)\phi C_{T_l} + q - \lambda_2 - \lambda_1(1-\alpha) \quad (23)$$

$$\frac{\partial W}{\partial T_H} = 0 \Leftrightarrow \lambda_3 = q(\phi C_{T_H} - 1) + \alpha + \lambda_2) \quad (24)$$

Note first that the citizen will always set the probability of re-election equal to 1 to maximize the level of public good provided. Trivially, equation (20) implies that $\frac{\partial W}{\partial \sigma_H} > 0$ and $\sigma_H = 1$. Combining equations (21) and (22) similarly gives $\frac{\partial W}{\partial \sigma_L} = Z(1-q) > 0$ and $\sigma_L = 1$. Equation (22) shows that one of the participation constraints must bind - if not, public good in one of the states could be increased whilst keeping taxes constant.

Suppose the participation constraint in the L case (constraint 1) does not bind. This implies that $\lambda_2 = 1$ so that the participation constraint in the H case binds and G_L is set such that $G_L = (\bar{F}(1+u) + T_L)(1-\alpha) + \alpha(T_L + T_H) + \sigma_L Z$. The participation constraint in the L case implies that the optimal tax levels must respect:

$$(\bar{F}(1+u) + T_L^*)(1-\alpha) + \alpha(T_L^* - T_H) + \sigma_L Z \leq (\bar{F}(1-u) + T_L^*)(1-\alpha) + \sigma_L Z \quad (25)$$

$$\Leftrightarrow \alpha(T_L^* - T_H) \leq -(1-\alpha)\bar{F}2u \leq 0 \tag{26}$$

Intuitively we must have $T_H^* > T_L^*$ to ensure that the politician in the L case does not find it profitable to pretend he is in the H case. However $\lambda_2 = 1$ and $\lambda_1 = 0$ implies that

$$\lambda_4 = (1-q)\phi C_{T_l} + q - 1 \Leftrightarrow T_L^* = h(1/\phi),\ \lambda_4 = 0 \tag{27}$$

and

$$\lambda_3 = q(\phi C_{T_H} - 1) + \alpha \tag{28}$$

$$\Leftrightarrow T_H^* = h(q-\alpha)/\phi q),\ \lambda_3 = 0,\ q > \alpha \text{ or } \lambda_3 > 0, T_H^* = 0 \tag{29}$$

this implies that $T_H^* < T_L^*$ and violates (26). This completes the proof

A.2 Proof of propositions 1-3

The program lowers ϕ by $d\phi < 0$ and therefore increases taxes in both states :

$$\frac{\partial T_H^*}{\partial \phi} = -\frac{h'(1/\phi)}{\phi^2} < 0 \tag{30}$$

and

$$\frac{\partial T_L^*}{\partial \phi} = \min\{0, -\frac{h'(1-q-\alpha/\phi(1-q))}{\phi^2}\frac{1-q-\alpha}{1-q}\} \tag{31}$$

I write

$$\frac{\partial E(T^*)}{\partial \phi} = -\omega_1 < 0 \tag{32}$$

It also increases the spread between T_H^* and T_L^{*33} :

$$\frac{\partial T_H^*}{\partial \phi} - \frac{\partial T_L^*}{\partial \phi} = -\frac{h'(1/\phi)}{\phi^2} - (1-\alpha)(E(T^*) + \bar{F}) - \bar{F}u(1-\alpha) = -\omega_2 < 0 \tag{33}$$

From equations (30), (31) and (33) we can write the increase in average public good provision thanks to the program:

$$-\frac{\partial E(G^*)}{\partial \phi} = (1-\alpha)\omega_1 + (1-q)\alpha\omega_2 > (1-\alpha)\omega_1 \tag{34}$$

Consider an increase in $\bar{F}$ of the same amount ω_1. It leads to an increase in average public good provision such that:

$$\frac{\partial E(G^*)}{\partial \bar{F}}\omega_1 = (1-\alpha)(1-u)\omega_1 < -\frac{\partial E(G^*)}{\partial \phi} \tag{35}$$

This completes the proof of proposition 1.

---

[33]Provided $h(.)$ is not too concave.

Proposition 2 follows from observing that equation (35) implies that:

$$\frac{\partial E(S^*)}{\partial \bar{F}}\omega_1 > \frac{\partial E(S^*)}{\partial \phi} \tag{36}$$

as $E(S^*) = E(T^*) + E(F) - E(G^*)$. This gives proposition 2.

Proposition 3 follows from observing that as u increases $\frac{\partial E(G^*)}{\partial \bar{F}}\omega_1$ increases but $\frac{\partial E(G^*)}{\partial \phi}$ is unchanged.

A.3 Imperfect information on both sources of revenue

In this section I consider an alternative model in which the relative accountability advantage of taxes is obtained under the weaker assumption that tax revenues are better observed than transfers, but still imperfectly observed. To keep the model simple I assume that taxes are set by an incumbent who is limited by tax capacity $\bar{T}$ such that $T < \bar{T}$. The incumbent is maximizing rents, so he sets $T = \bar{T}$ (we will see below that rents are increasing in T). The structure of this model is otherwise as above.

Asymmetries of information are introduced by assuming that the citizen observes $\hat{T}$ and $\hat{F}$, which are related to real revenues by $T = \hat{T}(1 + v)$ and $F = \hat{F}(1 + v)$, where $v < 1$. She observes real revenues with some probability p_T for taxes and p_F for transfers. The assumption that tax revenues are better observed than transfer revenues is captured by the fact that $p_T > p_F$.

Assume for now that $v > 0$: citizens systematically under-estimate the level of revenues. It is always optimal for the citizen to set the probability of re-election σ equal to 1 if the politician provides $G = (\hat{T} + \hat{F})(1 - \alpha) + Z$ when she receives no signal, $G = (\hat{T} + \hat{F}(1 + v))(1 - \alpha) + Z$ when she observes the true F, $G = (\hat{T}(1 + v) + \hat{F})(1 - \alpha) + Z$ when she observes the true T and $G = (\hat{T} + \hat{F})(1 - \alpha)(1 + v) + Z$ when she observes all revenues. The expected value of the public good is thus:

$$E(G) = (\hat{T}(1 + p_T v) + \hat{F}(1 + p_F v))(1 - \alpha) + Z \tag{37}$$

And expected rents are:

$$E(S) = \alpha\{\hat{T}(1 + p_T v) + \hat{F}(1 + p_F v)\} + v(\bar{T}(1 - p_T) + \bar{F}(1 - p_F)) - Z \tag{38}$$

Rewriting, we get the expected share of rents in total revenues as a function of the share of transfers in total revenues $E(f)$:

$$E(s) = \alpha + Z/R + \frac{v(1 - p_T)}{1 - u} + E(f)v\frac{p_T - p_F}{1 - v} \tag{39}$$

In this context it is straightforward to see that an increase in tax revenues leads to a bigger increase in the public good than an increase in transfers of the same amount, and that an increase in taxes (transfers) lowers (increases) the share of rents in total revenues. We can think of the tax policy studied here as an increase in $\bar{T}$ and hence T and of the transfer policy as an increase in F.

If we instead assume that $v < 0$, ie citizens systematically over-estimate the government's revenues, there is no relative accountability advantage of taxes. To see this note that as long as

43

the citizen does not observe revenues perfectly her mis-perception of the politician's participation constraint leads her to offer a menu $\{G, \sigma\}$ which asks the politician to provide more public good than the maximum he is willing to give in exchange for re-election. The politician will therefore choose to divert maximal rents and forgoe re-election (setting $G = (1 - \alpha)R$) with probability $1 - p_T p_F$, but will provide $G = (1 - \alpha)R + Z$ when the citizen perfectly observes public revenues (probability $p_T p_F$).

B   Transfer allocation appendix

B.1   The transfer allocation rule

The most important source of municipal revenue is the *Fundo de Participacao dos Municipios* (FPM), an automatic federal transfer established by the Brazilian Constitution. The FPM allocation mechanism divides local governments into population brackets which determine the share of their state's total FPM resources they will receive. Smaller population brackets are allocated lower shares. Each of the 26 states receives a different share of the total FPM resources in the federal budget, so two municipalities will receive the same amount only if they are in the same population bracket and state. The revenue sharing mechanism determining the amount $FPM_{i,t}^s$ received by government i in state s is

$$FPM_{i,t}^s = \frac{f(pop_{i,t})}{\sum_{j \in s} f(pop_{j,t})} FPM^s \tag{40}$$

where $f(pop_{i,t})$ is the coefficient corresponding to the population bracket in which the local government's population is found. Table 9 presents the population brackets and associated coefficients[34] in its first two columns.

Why would federal politicians tie their hands and allocated resources based on objective criteria rather than use discretion? Litschig (2011) shows it can be explained by the political agenda of the military dictatorship that put it in place in 1967. One of the objectives of the military was to wrest control of resources away from the traditional political elite and their clientelistic pratices and depoliticize public service provision. Creating a revenue sharing fund for the municipalities based on a objective criteria of need served this agenda. The choice of population brackets limits incentives for local officials to try to tinker with their population figures compared to a linear schedule. The initial legislation created cutoffs at round population numbers (multiples of 2000, 4000, 30000, etc) and stipulated that these should be updated regularly to keep up with population growth at the national level. This updating probably explains the value of the exact cutoffs we observe today - though they are never updated during our period of interest. The technical, a-political approach of the military in designing the rule is still reflected in the fact that the thresholds are still equidistant from one another today.

This rule has already been used by several authors (see Brollo *et al.* (2012), Litschig and Morrison (2010), Litschig (2011)), exploiting it for identification purposes is not a novelty of this paper. This appendix does not attempt to provide an exhaustive discussion of the merits of this identification strategy - the interested reader is referred to the previous papers on the subject for a more thorough discussion. It merely presents important checks for the validity of this strategy.

The *Tribunal de Contas Uniao* (TCU) determines how much each municipality will receive each year using the population estimates calculated by the Brazilian Statistical Institute (IBGE). The IBGE uses a top-down approach to produce these estimates. In a nutshell IBGE first produces a population estimates for a large area, say the whole of Brazil, based on estimated birth and mortality rates and migration. It then allocates this population to Brazilian states based on their share of total population in the last Census and their growth rate between the last two Census. The same logic applies to the allocation of population within states to municipalities (see Brollo

---

[34]Set by Decree No. 1881/81 and unchanged since 1981.

Table 9: Real and theoretical *FPM* transfers per capita and coefficients

Population	Coefficient	Real Transfer	Theoretical Transfer
<10,189	0.6	385	341
10,189-13,584	0.8	192	170
13,585-16,980	1	175	166
16,981-23,772	1.2	160	47
23,773-30,564	1.4	142	54.6
30,565-37,356	1.6	131	63.7
37,357-44,148	1.8	122	71.7
44,149-50,940	2	117	80.6
50,941-61,128	2.2	108	87.8
61,129-71,316	2.4	99	94.8
71,317-81,504	2.6	92	103.7
81,505-91,692	2.8	85	107.5
91,693-101,880	3	84	125.6
101,881-115,464	82	127.8	134.8
115,465-129,047	76	132.6	135.3
129,048-142,632	68	144.4	146.1

Population is the official population estimate from the IBGE. The coefficient are obtained from official documents of the *Tribunal de Contas Uniao* and used to estimate the theoretical FPM transfer allocated to each municipality. Real FPM transfers received are published by the *Tesouro Nacional*.

et al. (2012) for more details). The IBGE is an independent agency which does not interact with local governments and prides itself on its political independence. If municipalities wished to tinker with their estimated population numbers it is unlikely that they would choose to try to influence the IBGE. The TCU bases its allocation of transfer revenues on the estimates of the IBGE but does not always perfectly follow the rule. During the 1990s several municipalities split and this reduced the population size of pre-existing municipalities. Several low amendments froze the FPM coefficients to avoid large decreases in FPM transfers. The federal government ruled that by 2008 all municipalities should received the amount of FPM transfers corresponding to their actual population estimates, and established a transition period to the new regime. This explains why the observed FPM transfers do not perfectly follow the rule. What's more the population figures used by TCU (published each year) do not always perfectly coincide with the estimates by the IBGE, suggesting some manipulation may occur at this level (Litschig and Zamboni (2008) finds some evidence of manipulative sorting in 1989 and 1991 in the TCU figures). We therefore expect any manipulative sorting to occur in the TCU data, not the IBGE estimates. I use the IBGE estimates of population size in all the analysis of this paper.

I construct the amounts of theoretical FPM grants each municipality is allocated according to the above rule depending on its state and population size thus estimated for each year. Table 9 reports the average of those theoretical grants as well as the average actual grants received by municipalities in each population bracket.It is clear from the table that population state and year do not perfectly predict the FPM grants each municipalities receives, due to several reasons.

46

Nonetheless, real FPM grants received do increase substantially at each population threshold.

B.2 Validity checks

Figure 6: Population density histogram
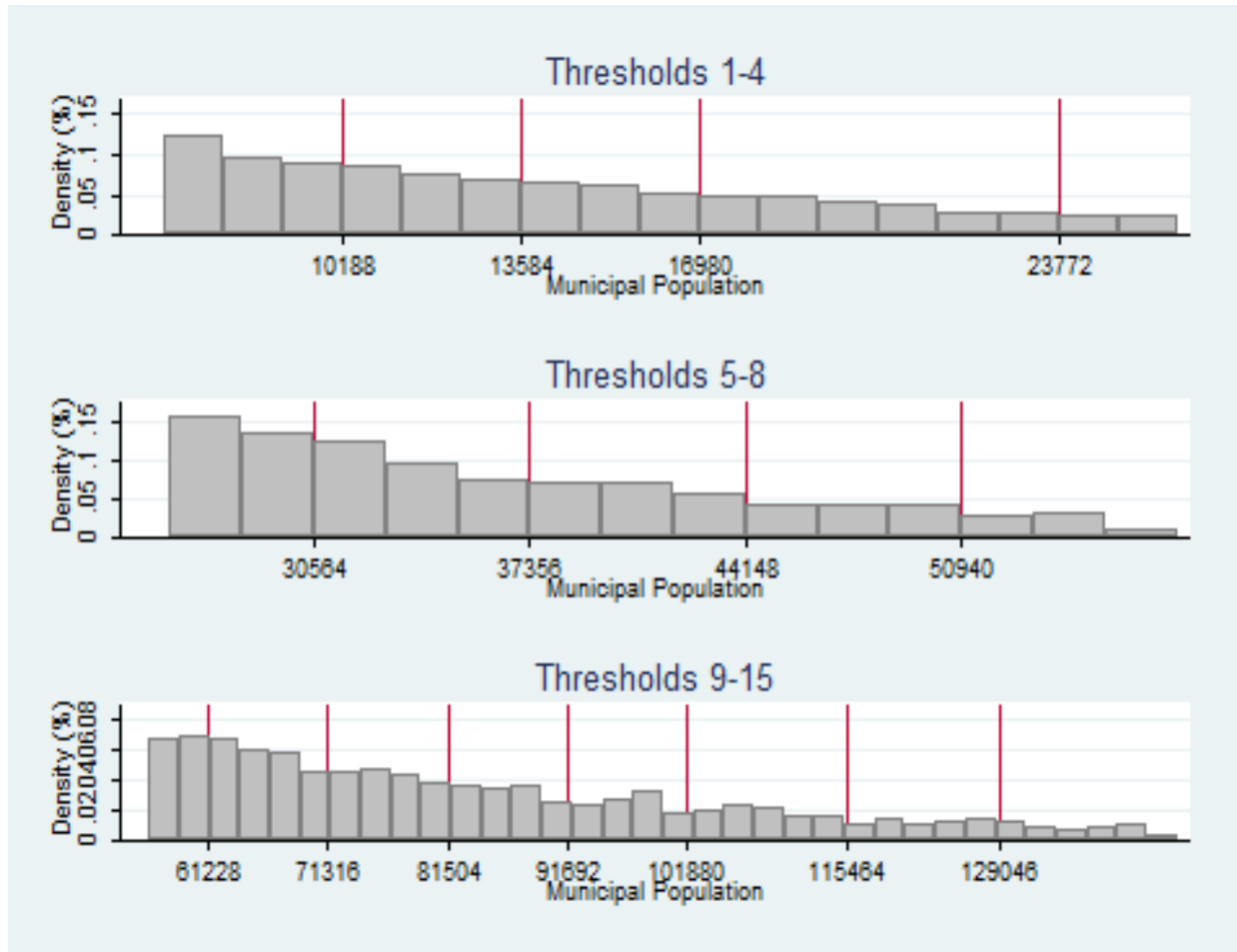
Bin sizes: 1132(top), 2264(middle), 2572(bottom).The bin sizes are chosen so that no bin contains a threshold.

Figure 7: McCrary density tests: pooled Thresholds
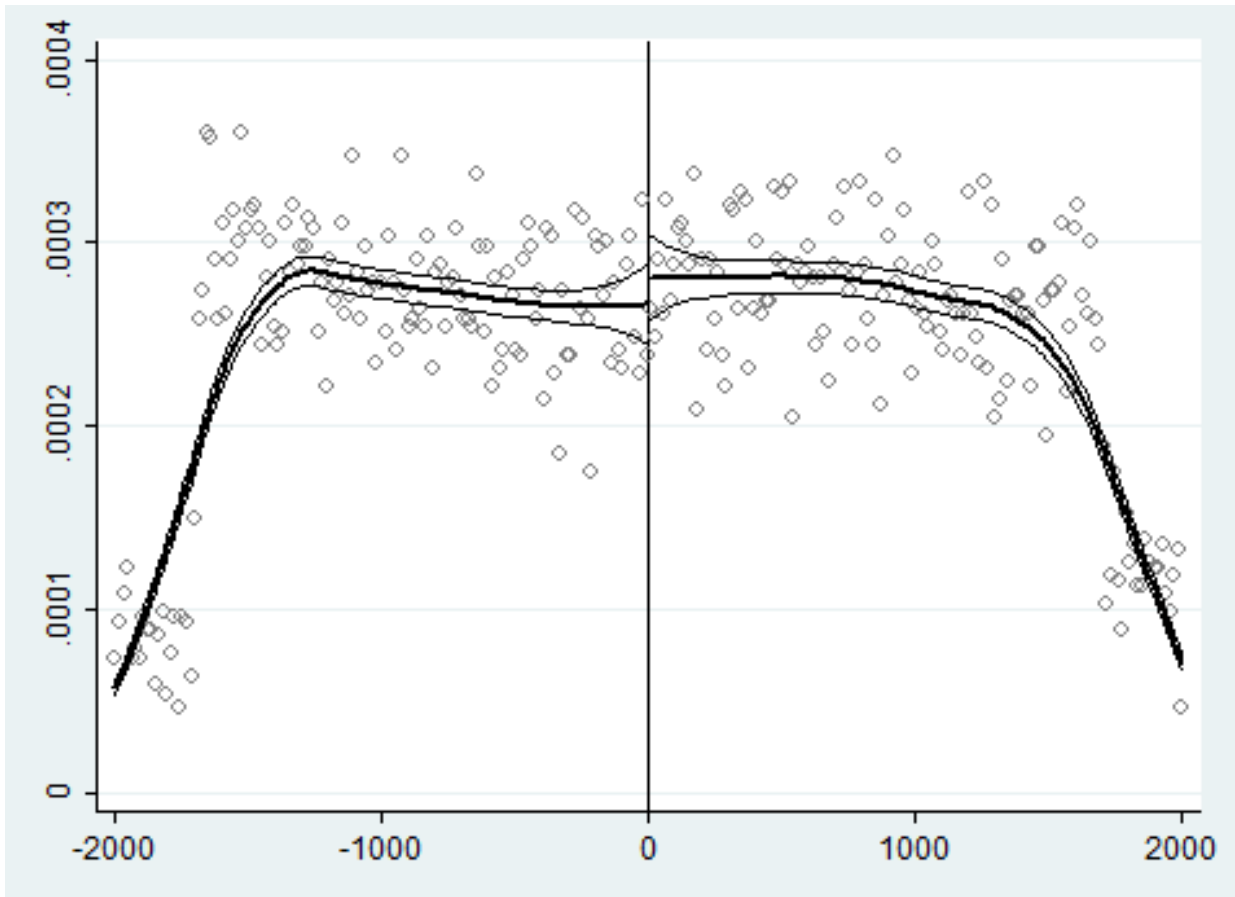
 Weighted kernel estimation of the log density according to population size performed separately on either side of the pooled threshold. Optimal binwidth and binsize as in McCrary (2008).

Figure 8: McCrary density tests: 15 thresholds separately

From top right to bottom left: thresholds 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15. Weighted kernel estimation of the log density according to population size performed separately on either side of the threshold. Optimal binwidth and binsize as in McCrary (2008).

Figure 9: The population discontinuity and population growth

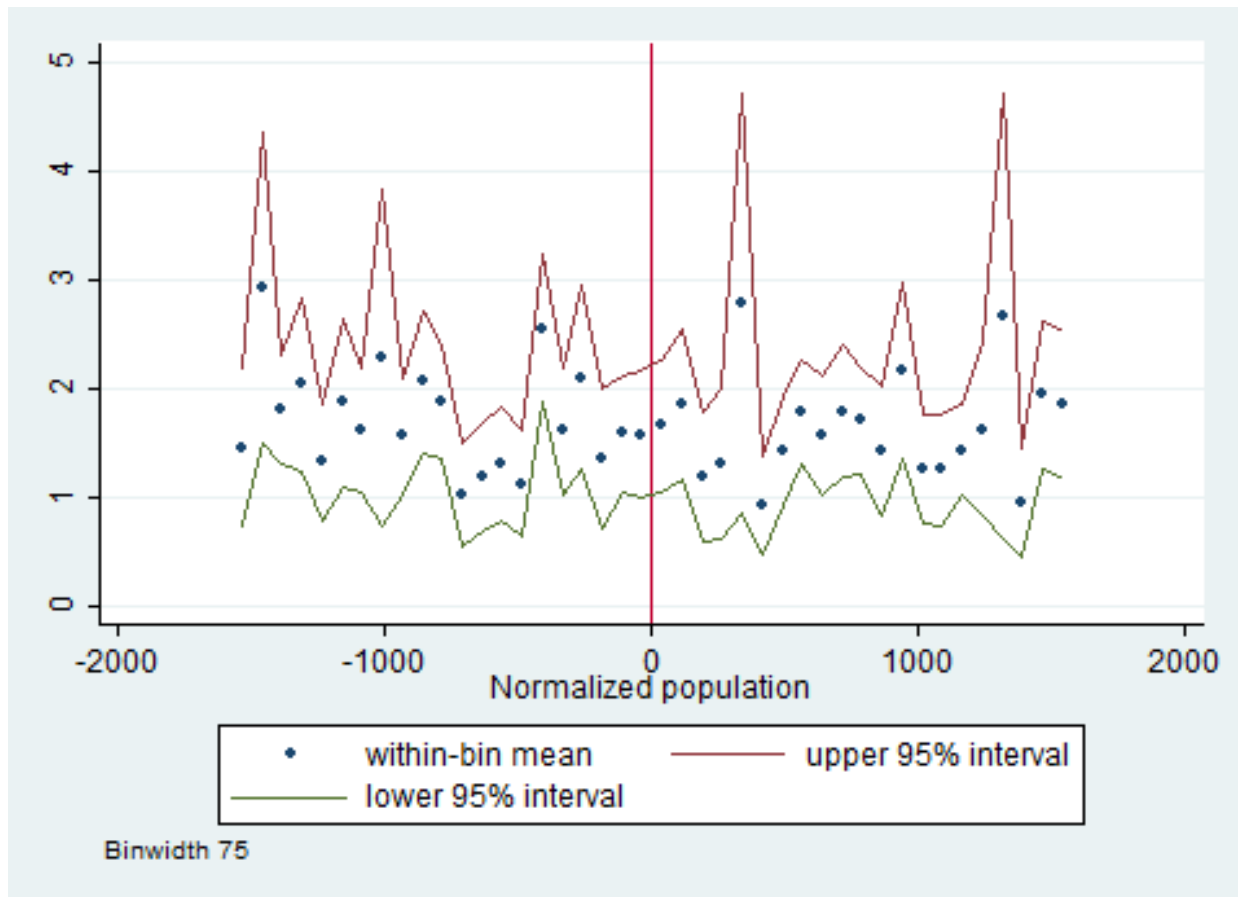The graph presents average population growth between year t and $t+1$ averaged over 75 inhabitants bins over normalized population size in year t.

Table 10: Descriptive statistics of main variables by FPM threshold

	Threshold 1	Threshold 2	Threshold 3	Threshold 4	Threshold 5	Threshold 6	Threshold 7
Taxes per capita 1998	47.95	42.51	48.26	53.67	60.29	71.76	73.96
	(129.2)	(54.95)	(95.75)	(84.78)	(106.0)	(122.3)	(175.9)
School infrastructure 1998	4.208	4.161	3.997	3.685	3.530	3.106	3.283
	(2.157)	(1.829)	(1.865)	(1.775)	(1.644)	(1.539)	(1.635)
School quality 1998	-0.735	-0.721	-0.823	-0.860	-0.962	-0.761	-0.947
	(1.315)	(1.334)	(1.262)	(1.241)	(1.224)	(1.190)	(1.143)
Health infrastructre 1999	3.649	4.521	5.337	6.571	7.907	9.377	10.16
	(2.515)	(3.028)	(3.596)	(4.437)	(4.746)	(5.293)	(5.453)
Public revenues per capita 1998	423.8	394.4	387.6	376.4	359.8	354.9	343.3
	(217.3)	(192.3)	(175.5)	(184.1)	(156.2)	(155.7)	(211.0)
Population 1998	9784.2	13026.1	16571.0	21994.7	28152.6	34433.8	39909.2
	(1546.1)	(1788.3)	(2171.0)	(3009.9)	(3387.0)	(4014.4)	(4809.4)
GDP per capita in 1999	4543.0	4466.2	4789.6	4807.3	4881.0	4893.3	4585.2
	(3697.0)	(3897.8)	(3964.4)	(3668.5)	(3967.9)	(3661.4)	(3100.7)
Weight	0.359	1.569	0.973	0.834	1.902	1.356	1.597
	(4.472)	(18.27)	(10.77)	(1.181)	(2.332)	(1.419)	(1.582)

	Threshold 8	Threshold 9	Threshold 10	Threshold 11	Threshold 12	Threshold 13	Threshold 14	Threshold 15
Taxes per capita 1998	67.04	72.71	86.56	95.16	89.86	106.5	107.2	119.8
	(84.69)	(95.80)	(98.65)	(100.3)	(90.32)	(86.30)	(98.41)	(133.7)
School infrastructure 1998	3.210	3.092	2.772	2.435	2.139	2.197	2.244	2.034
	(1.514)	(1.477)	(1.206)	(1.031)	(0.889)	(1.125)	(1.099)	(0.754)
School quality 1998	-0.797	-0.725	-0.515	-0.202	-0.356	-0.162	-0.253	-0.496
	(1.102)	(1.198)	(1.170)	(1.204)	(1.058)	(1.136)	(1.098)	(1.054)
Health infrastructre 1999	12.68	12.88	16.53	16.03	14.95	14.15	16.38	18.00
	(7.090)	(7.011)	(11.41)	(11.57)	(8.536)	(5.527)	(6.849)	(7.902)
Public revenues per capita 1998	353.8	346.8	366.3	370.8	340.7	394.4	396.7	391.4
	(294.8)	(238.3)	(190.7)	(197.6)	(146.3)	(205.2)	(262.0)	(301.1)
Population 1998	47148.7	56082.6	64160.5	74253.9	82325.6	90958.0	103026.7	117479.1
	(5206.8)	(6355.7)	(6345.1)	(7864.3)	(9333.5)	(10620.9)	(11259.8)	(12194.3)
GDP per capita in 1999	5414.9	5291.1	6604.1	7001.5	5720.3	7113.5	7272.6	7126.0
	(6413.6)	(5275.1)	(5468.8)	(5916.4)	(3546.8)	(4738.3)	(6093.8)	(5721.6)
Weight	2.003	1.828	2.438	1.421	0.641	0.733	0.698	0.681
	(1.731)	(1.668)	(1.835)	(1.569)	(0.557)	(0.423)	(0.447)	(0.590)

Mean (standard deviation). Each cell contains the average value of the row variable in municipalities around the column threshold. The interval around a threshold is plus or minus 1698 inhabitants for thresholds 1 to 3, plus or minus 3396 for thresholds 4 to 8, plus or minus 5094 for thresholds 9 to 13 and plus or minus 6792 for thresholds 14 and 15. Each interval is constructed so that each municipality with population size between 8490 and 143'633 is in one and one only interval around a threshold.

Table 11: Marginal propensity to spend out transfer revenues by threshold

	Quant. education infra.		Qual. education infra.		Health infra.	
Threshold 1						
Transfers per capita	0.098	(0.062)	0.116*	(0.065)	-0.141*	(0.072)
Observations	3682		3680		1205	
Threshold 2						
Transfers per capita	0.106**	(0.500)	0.161**	(0.080)	-0.153	(0.171)
Observations	2982		2982		914	
Threshold 3						
Transfers per capita	0.077	(0.052)	0.027	(0.033)	0.083	(0.168)
Observations	3250		3249		1047	
Threshold 4						
Transfers per capita	0.093***	(0.029)	0.093**	(0.047)	-0.222	(0.189)
Observations	2796		2790		881	
Threshold 5						
Transfers per capita	0.049	(0.083)	0.020	(0.077)	-0.083	(0.634)
Observations	1966		1961		614	
Threshold 6						
Transfers per capita	-0.051	(0.114)	0.134	(0.168)	-0.281	(1.076)
Observations	1146		1146		325	
Threshold 7						
Transfers per capita	0.145	(0.105)	0.112**	(0.062)	-1.280	(0.818)
Observations	852		849		241	
Threshold 8						
Transfers per capita	0.055	(0.079)	0.071	(0.124)	-0.441	(0.925)
Observations	698		697		194	
Threshold 9						
Transfers per capita	0.067	(0.081)	0.130*	(0.075)	0.269	(1.417)
Observations	715		715		206	
Threshold 10						
Transfers per capita	0.099*	(0.055)	-0.156	(0.208)	2.412	(4.399)
Observations	540		540		146	
Threshold 11						
Transfers per capita	0.078	(0.083)	0.110	(0.194)	1.025	(1.445)
Observations	414		414		106	
Threshold 12						
Transfers per capita	-0.156	(0.232)	0.159*	(0.086)	-1.576	(5.726)
Observations	328		328		87	
Threshold 13						
Transfers per capita	0.016	(0.119)	0.069	(0.235)	-4.072	(2.836)
Observations	286		286		52	
Threshold 14						
Transfers per capita	-0.127	(0.119)	-0.051	(0.223)	-2.764	(5.649)
Observations	230		230		55	
Threshold 15						
Transfers per capita	0.110	(0.171)	0.122	(0.147)	-2.866	(3.394)
Observations	236		236		62	

C Propensity score appendix

The propensity score used to implement the weighted-difference in differences methodology is calculated by estimating a probit model of the probability that a municipality started a PMAT program sometime between 1998 and 2008 as a function of the pre-intervention characteristics used in the first column of Table 3. Table 12 presents the results of this estimation. This model is then used to predict the propensity (probability) that a municipality will privatize.

Table 12: Determinants of the probability of joining a program

Income	0.1586**
	(0.0757)
Population	0.1069**
	(0.0468)
Taxes in 1998	0.0000
	(0.0005)
Agr\ GDP	-0.0048
	(0.0042)
Serv\ GDP	0.0006
	(0.0045)
Education	0.0203
	(0.0875)
Urban pop.	0.8583***
	(0.2942)
Inequality	-0.8831
	(0.9148)
Governor's party	-0.0827
	(0.1073)
Pol. competition	0.7736**
	(0.3786)
Observations	3560

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient represents a marginal effect and the regression includes state fixed effects.

I identify control and treatment observations on a common support as follows. I exclude all control observations whose propensity scores are less than the propensity score of the treated

municipality at the mid-point of the first percentile of the treatment propensity score distribution, and exclude all treated observations whose propensity score is greater than the propensity of the control observation at the mid-point of the 99th percentile of the control distribution. This eliminates 33% of control observations and 10% of treated. Figure 10 graphs the distribution of the propensity score in the treated and control groups. The red lines indicate the limit of the common support.
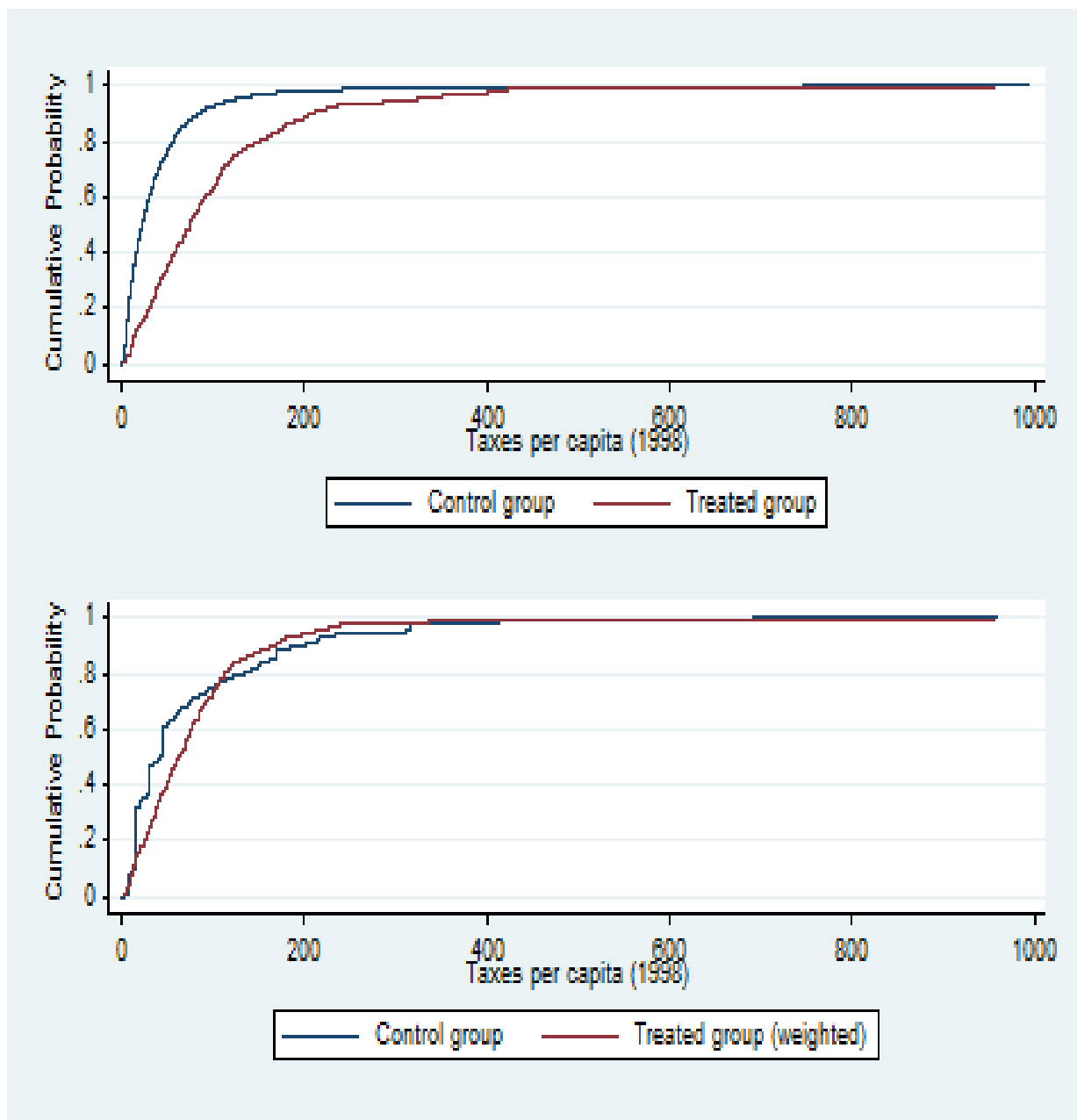
Figure 10: Distribution of the propensity score
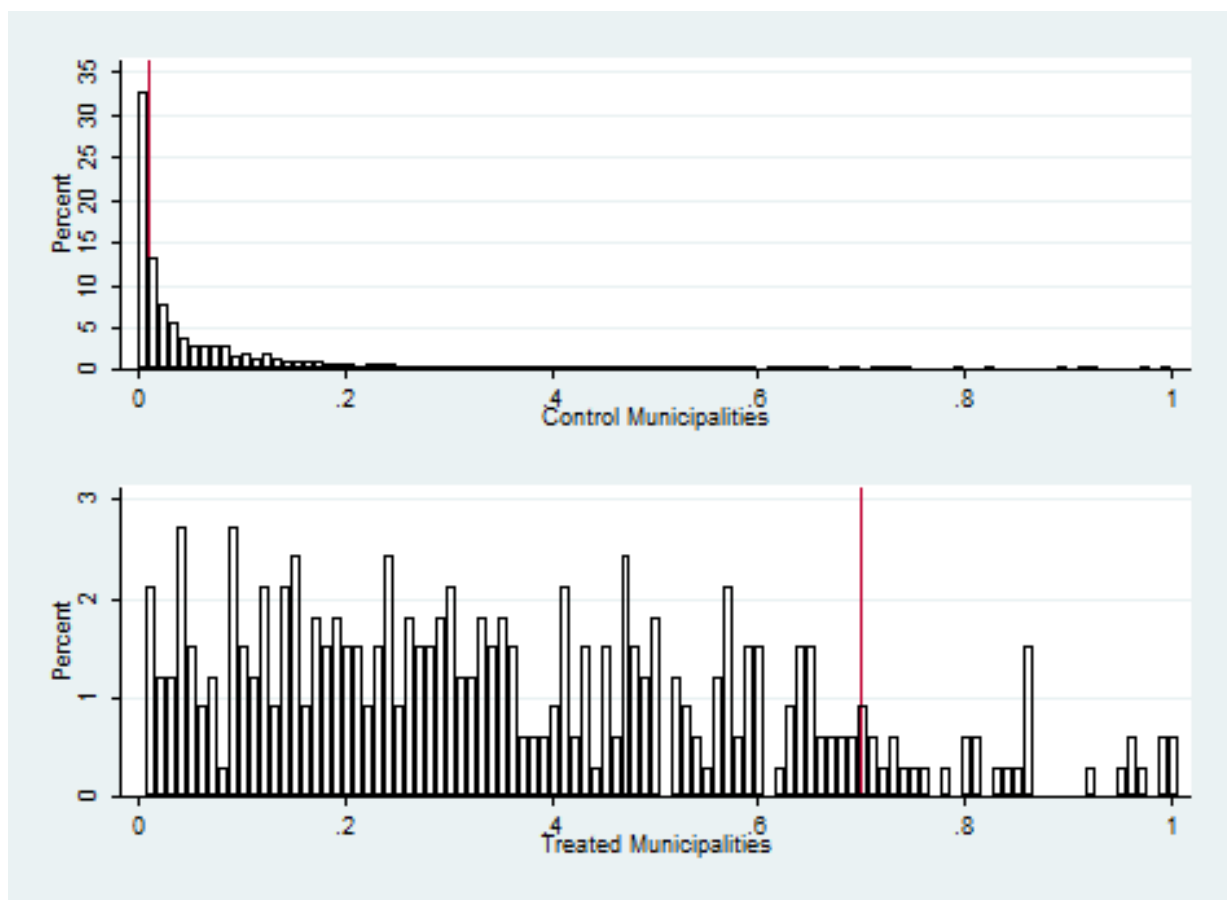
 Density distribution of the estimated probability of joining the program. Red lines indicate the limits of the common support sample.
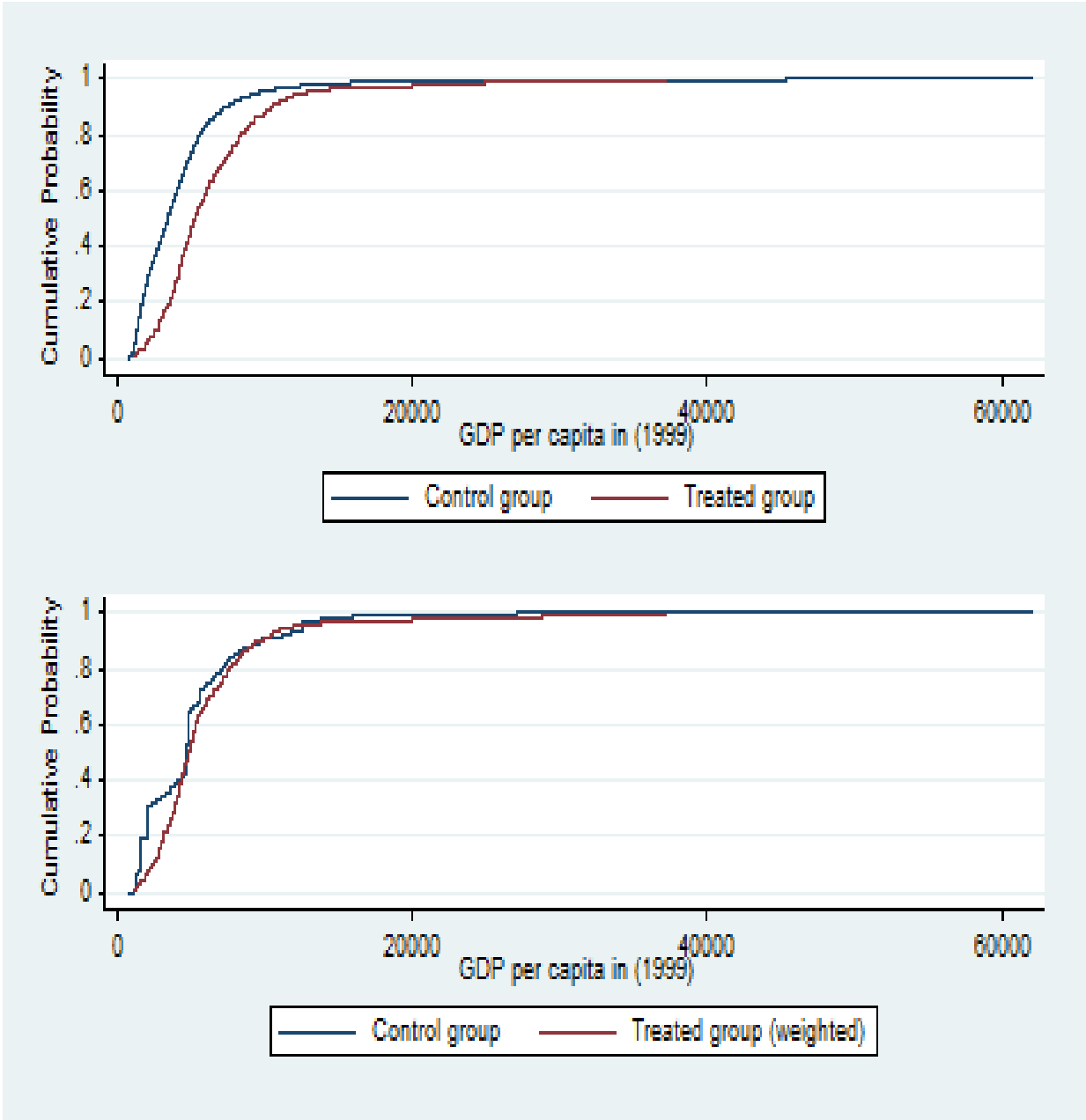
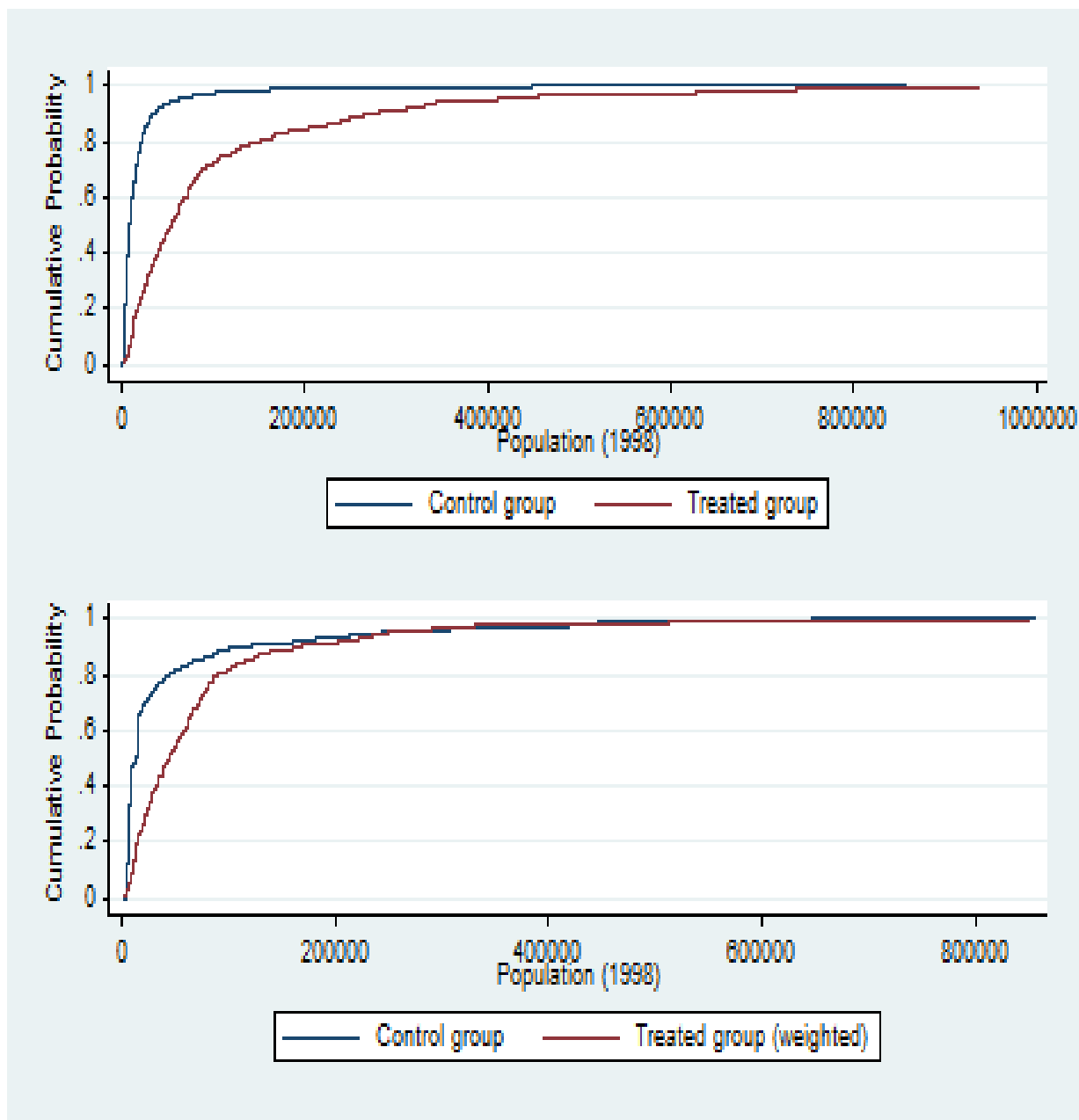Figure 11: CDF of tax revenues per capita in 1998 in treated and control municipalities

 The graphs plot the CDF of tax revenues per capita in 1998 in treated and control municipalities. The top panel plots the distributions for the entire sample. The bottom panel plots the distribution for the common support sample, the control municipalities are weighted by a function of their estimated probability of joining the program.

Figure 12: CDF of GDP per capita in 1999 in treated and control municipalities

 The graphs plot the CDF of GDP per capita in 1999 in treated and control municipalities. The top panel plots the distributions for the entire sample. The bottom panel plots the distribution for the common support sample, the control municipalities are weighted by a function of their estimated probability of joining the program.

Figure 13: CDF of population in 1998 in treated and control municipalities

The graphs plot the population in 1998 in treated and control municipalities. The top panel plots the distributions for the entire sample. The bottom panel plots the distribution for the common support sample, the control municipalities are weighted by a function of their estimated probability of joining the program.

D Table appendix

Table 13: Impact of the program, alternative specifications

	Taxes		Educ infrastructure: quantity		Health infrastructure		Corruption	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
	(1): All sample, (2): Logs							
All years	12.581***	0.098***	0.249***	0.046***	1.3	0.071**	-53.248	-0.135
	(2.235)	(0.023)	(0.038)	(0.022)	(0.791)	0.036	(41.890)	(0.122)
Observations	54577	28112	54589	28215	17121	9999	688	526

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All regressions include municipality and year fixed effects as well as controls for GDP per capita, population size, share of agriculture and services in GDP, and changes in political competition, mayor's party affiliation and mayor's term limit in the previous election. The results for the log specification (2) are obtained from running propensity-score weighted versions of equation (14) on the common support sample using the natural logarithm of taxes, infrastructure, enrollment and the corruption index as dependent variables.

Table 14: Impact of the program by time between application and program start

Time between application and program start	0 year	1 year	2-3 years
3 years before	4.298	-0.540	-1.149
	(3.314)	(2.782)	(2.367)
2 years before	-1.994	-2.675	-0.955
	(3.572)	(3.210)	(3.269)
1 year before	0.361	0.994	1.557
	(4.049)	(3.918)	(3.415)
Program : all years	8.019*	10.248**	10.251***
	(4.581)	(4.634)	(3.816)
Observations	25436	26584	25593
Clusters	2374	2480	2389

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables is real tax revenues per capita. All regressions include municipality and year fixed effects and controls for GDP per capita, population size, share of agriculture and services in GDP, political competition in the previous election, mayor's party affiliation and whether the mayor is a facing a term limit.

Table 15: Reduced form impact of the program on spending outcomes

	Educ. infrastructure: quantity		Educ. infrastructure:quality		Health infrastructure		Corruption
s(1 ≤ j ≤ 11)	0.222*** (0.079)		0.224*** (0.062)		1.137* (0.647)	1.209 (1.367)	-9.114* (5.205)
before (j = −4)		-0.070 (0.156)		0.089 (0.093)			
before (j = −3)		-0.043 (0.182)		0.054 (0.059)			
before (j = −2)		0.017 (0.102)		0.046 (0.055)			
(j = 1)		0.163 (0.121)		0.227** (0.090)			
(j = 2)		0.161 (0.117)		0.197* (0.111)			
(j = 3)		0.220 (0.140)		0.263** (0.117)			
(j = 4)		0.282** (0.135)		0.280*** (0.103)			
(j = 5)		0.335** (0.199)		0.258*** (0.094)			
$PMAT$ (−10 ≤ j ≤ 2)						0.059 (1.294)	
eated (−10 ≤ j ≤ 11)						11.452 (8.499)	
tions	28215	26899	28182	26866	9999	9999	526

bust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include year fixed effects and time-varying controls (GDP per capita, population size, the griculture and services in GDP, whether the mayor is in his second term, and political competition in the last election). Municipal fixed effects are included in all but the last column. ons without fixed effects (last column) include state fixed effects and as controls a indicator of whether the municipality joins a $PMAT$ program between 1998 and 2008 and municipal tics (income per capita, inequality, median education, life expectancy, population density, all in 2000, and whether the municipality has a local radio station in 1998 and a seat of the ranch in 1998). The quantity of education infrastructure is the number of classrooms in use in municipal schools per 1000 inhabitants, the quality of education infrastructure is a municipal lity index constructed as explained above, health infrastructure is the number of municipal health establishments per 100,000 inhabitants, and corruption is the corruption index described

Table 16: Impact of the program on expenditure outcomes by time between application and program start

Time between application and program start (years)	Educ. infrastructure: quantity			Educ. infrastructure:quality		
	0	1	2-3	0	1	2-3
3 years before	0.023	-0.016	-0.015	0.022	-0.025	-0.091
	(0.203)	(0.099)	(0.165)	(0.150)	(0.081)	(0.114)
2 years before	0.102	0.010	0.038	-0.014	-0.040	-0.045
	(0.212)	(0.115)	(0.176)	(0.146)	(0.083)	(0.147)
1 year before	0.021	0.012	-0.006	0.020	0.010	0.021
	(0.221)	(0.116)	(0.171)	(0.131)	(0.098)	(0.139)
Program : all years	0.217	0.210*	0.218	0.230**	0.200*	0.199
	(0.202)	(0.121)	(0.152)	(0.125)	(0.110)	(0.164)
Observations	25535	26686	25692	25503	26657	25664

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The quantity of education infrastructure is the number of classrooms in use in municipal schools per 1000 inhabitants, the quality of education infrastructure is a municipal school quality index constructed as explained above and health infrastructure is the number of municipal health establishments per 100,000 inhabitants. All regressions include municipality and year fixed effects and controls for GDP per capita, population size, share of agriculture and services in GDP, political competition in the previous election, mayor's party affiliation and whether the mayor is a facing a term limit.

Table 17: Impact of the program on GDP and population

	GDP	Population
Program : all years	0.742	-0.094
	(0.902)	(0.127)
Observations	24070	24070
Municipalities	2462	2462

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All regressions include year fixed effects, controls are GDP per capita, population size, share of agriculture and services in GDP, and changes in political competition, mayor's party affiliation and mayor's term limit in the previous election.

Table 18: Within municipality means (SD) of taxes and transfers

	All	Controls	Treated before $PMAT$	Treated after $PMAT$
Taxes	70.1 (74)	65.1 (78)	90 (19)	142.9 (25)
Transfers	174.8 (48)	182 (77)	106 (17)	129 (22)