Tax Me, But Spend Wisely : The Political Economy of Taxes, Theory and Evidence from Brazilian Local Governments

Lucie Gadenne*

October 2011

Abstract

This paper shows that local governments are more accountable when they have more capacity to tax their citizens. I construct a principal-agent model of public finance in which public revenues come from taxes and inter-governmental transfers. An increase in taxes keeping transfers constant 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 consider a program in Brazil that invests in the modernization of local tax administrations. Using 10 years of panel data and quasi-exogenous variations in the timing of program uptake I find that the program increases tax collection of local governments by 11% after four years. This increase in taxes is used to raise local public good provision but not corruption : the share of resources diverted by local politicians in total public revenues decreases as taxes increase. A discontinuity in the rule allocating federal transfers to local governments is used to compare the impact on spending outcomes of an exogenous increase in transfers to that of a raise in tax revenues thanks to the program. Results show, in line with the model's predictions, that local officials use extra tax revenues more to increase the supply of municipal public goods and less on private rents than they do with extra transfer revenues.

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

*PhD Candidate, Paris School of Economics. Email : gadenne@pse.ens.fr. I am grateful to Denis Cogneau, Claudio Ferraz and Thomas Piketty for their advice and support throughout the project. I would also like to thank Luc Behaghel, Tim Besley, Clement de Chaisemartin, Julien Grenet, Walker Hanlon, Karen Macours, Guy Michaels, Tommaso Nannicini, Ernesto Stein, Eric Verhoogen and Ekaterina Zhuravskaya for their helpful comments as well as Stephan Litschig and Yves Zamboni for sharing their data, Marcelo Fernandes and Rogerio Goulart for their help understanding the PMAT program and seminar participants at the Paris School of Economics, the London School of Economics, Bocconi University, the BNDES, PUC-Rio, FGV-Rio, the LACEA Political Economy Workshop 2011, the European Political Science Association Meeting 2011 and the IV Fiscal Federalism Workshop. I acknowledge financial support from the AXA Research Fund and the CEPREMAP. Part of the research was done while the author was visiting the London School of Economics as an AMID fellow.

1 Introduction

Local governments around the world generate very little of the revenue they spend. Scholars of local governments in developing countries also point out the poor quality of their governance and its impact on local public spending, ranging from carelessness in public good provision to outright diversion of public resources (Rose-Ackerman (1999)). Yet decentralization has been pushed forward for the last 30 years around the developing world as a way of bringing governments and service delivery closer to the people, with mixed results (see Bardhan and Mookherjee (2006) for a review of the evidence). The idea that how these governments are financed and in particular how much taxes they collect may affect the quality of their public spending has recently been flagged out in the policy and development studies literature[1]. Does this mean that making local governments collect more taxes will improve how they spend their revenues - and if yes, what kind of policy can increase local tax collection? Or will any increase in local budgets just lead to more inefficient spending and revenue diversion by local politicians?

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 quality of their public spending. The more they rely on tax collection, the more local politicians have to respond to their constituents' demands with respect to public good provision and the less rents they can extract for their private use. This idea has been voiced before in policy debates - the World Bank for example argues that *"To increase responsiveness to local citizens, subnational governments need a local tax instrument and the freedom to set tax rates."* (World Bank (2004)) but is mostly supported by anecdotal evidence. This paper seeks to assess the validity of that statement theoretically and empirically. 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 amongst Brazilian local governments using a program that increases local tax collection, exogenous variation in federal transfers received by local governments, and panel data on public revenues, public good provision and corruption at the local government level.

This paper's theoretical framework consists in a political agency model of public finance in which public revenues come from local taxes that are endogenously determined and exogenous transfers. A rent-seeking incumbent politician decides how to allocate the public budget between public good provision that benefits citizens 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 the extent to which citizens can control the allocation of tax and transfer revenues : increasing the amount of tax revenues leads to a fall in the share of the total budget that is diverted by the politicians as information rents decrease. The model's

[1] See in particular Bird and Smart (2002), Brautigam et al. (2008) and Moore (2007) for a review of this literature.

key prediction is that a policy that increases the efficiency of the tax administration and thus increases local tax revenues will lead to a bigger rise in local public good provision, and less increase in revenue diversion for private use, than a policy that increases transfer revenues by the same amount. An increase in the government's capacity to tax its citizens thus makes it more accountable, in the sense that it leads to an allocation of public revenues towards more expenditures that benefit citizens at the expense of corruption expenditures which here only benefit the incumbent politician.

I then test the model's predictions by evaluating the impact of a federal program that modernizes the tax administration of local governments (municipalities) in Brazil : the program offers municipalities subsidized loans to invest in the efficiency of their local tax administration. Selection into the program is purely voluntary, I find that the 330 municipalities which join the program are substantially bigger and richer than the average municipality. The key empirical challenge this paper faces is thus that those governments' choice to participate may not be orthogonal to unobservable factors that also affect tax collection and/or the allocation of public revenues. The richness of the data and specificities of the program's institutional design however allow me to explore and rule out the most likely alternative explanations for the empirical results.

My empirical strategy relies on a difference in differences estimator and a unique dataset of 10 years of panel data on municipal tax and transfer revenues, quantity and quality of municipal education supply, corruption and wasteful spending by local officials from randomized audits of municipalities, and a large set of local level economic, demographic and political characteristics. The key characteristic of the program I use to check the plausibility of the identifying assumption 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 which create variations in the timing of program uptake that I argue are exogenous to unobserved local characteristics. I also present propensity-score weighted estimates that are robust to the existence of unbalanced pretreatment characteristics correlated with the dynamics of the outcome variable and restrict the sample to comparable municipalities.

I first consider how the program affects tax collection and then ask whether the increase in tax revenues benefited (corrupt) local politicians and/or citizens by evaluating whether the program led to an increase in corruption and/or more local education inputs. From this I compute Wald estimates of the impact of an increase in taxes thanks to the program on how much local politicians decide to allocate to education provision and corruption. I compare this estimate with the effect of an increase in transfers to test the model's prediction that an increase in taxes leads to better public spending outcomes than an increase in transfers. This effect is identified using a nonlinearity in the rule governing the allocation of the main federal transfer which leads to several exogenous discrete jumps in the amounts of transfers local governments receive, following the research design used by Brollo et al. (2010) and Litschig (2008a) which is not a novelty of this paper.

Results show that the tax modernization program raised local tax revenues by 11%

after four years and that this increase persisted over time. The cost of the investments in tax administration were on average recovered after only two years in the program. The raise in taxes was used to finance a 8% increase in municipal school infrastructure, an increase in school quality, but did not increase corruption. Comparing the impact of the increase in taxes thanks to the program to that of an increase in transfers of the same amount I find that an extra 10 Reais per capita of public revenues leads to at least 3 times more education inputs and local school enrollment when it comes from taxes than when it comes from federal transfers. Such an increase in transfer revenues leads to a nearly 30% raise in the occurrence of corrupt practices as a share of revenues whereas taxes lead to a decrease in corruption of roughly three times that amount, though this result is imprecisely estimated. These results suggest that local Brazilian governments spend increases in local taxes 'better' than increases in federal transfer revenues, but their interpretation is limited by the fact that the impact of taxes and of transfers is estimated on different sub-populations of Brazilian municipalities. I discuss to what extent these populations are comparable, and present some evidence that suggests that the mechanism outlined in the model - that citizens have better information on taxes than on transfers - partially explains the observed differences in how taxes and transfers are spent.

This paper's first contribution is a model in which an assumption which is fairly standard in the political economy literature (asymmetries of information) leads to the prediction that increases in tax capacity will make governments more accountable. The idea that relying on local taxes affects political officials' incentives dates 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) for a review). Zhuravskaya (2000) provides evidence for this mechanism amongst Russian cities. This paper differs by relying on an explicit political economy mechanism to explain why taxes lead to more public good spending than transfers; this mechanism will hold even if local governments cannot finance growth-enhancing local public goods[2]. The theoretical model developed here is also related to previous political agency models which argue that information asymmetries lead to more rent-taking opportunities by politicians (in particular Besley and Smart (2007)). 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 constituencies.[3].

The empirical result that tax capacity programs can increase local tax collection in Brazil contributes to the growing literature on state capacity and development which argues that governments' investment in their capacity to tax constitutes an important

[2]The type of public good provision this paper considers - education supply - 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. I provide some evidence that this theory cannot explain the results found in this paper.

[3]See Besley (2006) for a review of political agency models. One exception is Strumpf (1998) who builds on the idea that citizens may have different information on different sources of local public funds to explain the fly-paper effect

covariate of economic development (Besley and Persson (2009), Besley and Persson (2009)). To the best of my knowledge this paper is the first to consider specific innovations in tax collection and provide an empirical investigation of the returns to such innovations. Widely optimistic prognoses regarding the efficiency of investments in developing countries' tax administrations abound : OECD (2010b) for example reports the words of the President of the African Tax Administration Forum, Oupa Magashula, that investing in public resource mobilization can have up to "*a tenfold multiplier effect on states' resources*". I find an annual 'multiplier' effect of just over one (a one Real investment in tax administration leads to an extra one Real in tax revenue every year), far from tenfold but still very cost-effective.

Finally, this paper is also related to the large and growing literature on determinants of corruption and the quality of public expenditures at the local government level[4]. Of particular relevance here are recent papers on the impact on public expenditures or corruption of an increase in public resources amongst Brazilian local governments. One can easily take away from this literature the idea that a raise in local revenues leads to wasteful or corrupt government spending in Brazil[5] : Brollo et al. (2010) find that higher grants from the federal government leads to more corruption whilst both Caselli and Michaels (2009) and Ferraz and Monteiro (2010) show that windfall from oil royalties lead to no improvement in local public good provision but to an increase in public employment(Ferraz and Monteiro (2010)). This paper provides a framework that helps reconcile these previous findings with the idea that historically and cross-sectionally more government resources are associated with more efficient and accountable governments[6]. The model suggests that increases in public revenues coming from sources not directly observed by citizens, such as the ones considered in the literature, lead to worse public spending outcomes than increases in taxes, which generally constitute the bulk of government revenues; the research design allows for the first empirical comparison of the marginal increase of tax revenues and non tax revenues on the accountability of governments. This paper contributes more widely to the larger literature on the political economy of public good provision and corruption by focusing on the role of one institutional characteristic - governments' capacity to tax their citizens - as a determinant of accountability which has so far not been studied.

The outline of the paper is as follows. Section 2 provides an agency model of public finance that relates how governments are financed to how they allocate their budget between private rents and public good provision. Section 3 presents the institutional and economic context of the tax modernization program and explains why some municipalities choose to join the program whilst many do not. Section 4 evaluates the impact of the program on local tax and public spending outcomes and discusses the identifying assumption whilst section 5 contrasts the impact on education inputs and

[4]See for example Ferraz and Finan (2009) and Martinez-Bravo et al. (2011) on the role of elections, Olken (2005) and Litschig and Zamboni (2011) on bottom-up and top-down monitoring, or Reinikka and Svensson (2005) on the quality of the information available to users of public services.

[5]An exception is Litschig (2008a) who shows that more federal transfers leads to better education outcomes.

[6]See Lindert (2003).

corruption of higher taxes thanks to the program to that of an increase in transfers. I conclude with Section 6.

2 Model

2.1 Set-Up

Structure

The model follows the political agency framework of Besley and Smart (2007) in which a representative citizen decides whether to re-elect an incumbent politician without observing part of this politician's 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 a 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 (1_q) and $F_L = \bar{F}(1-u)$ in the low state L, where $u,\ q \in [0,1]$[7].

The incumbent politician maximizes the sum of rents extracted from being in office $S + \sigma Z$, where S is rents extracted from the budget, Z is the value of re-election and σ the probability of re-election. Rent-taking is limited by the budget constraint $T + F = R = G + S$, with G the level of public good, and institutional constraints which make it costly for the incumbent to run away with all public resources : maximal rent taking is $\bar{S} = \alpha R$ where $\alpha < 1$. I assume that capturing maximal rents is more attractive to the incumbent than getting no rents but being re-elected for sure, so $Z \leq \alpha(T+F), \forall T, F$[8]. 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 with $C(0) = C'(0) = 0$. I define $h(.) = C^{-1}(.)$.

Full information equilibrium

The citizen chooses for each state $i = H, L$ the reelection rule $\sigma(G_i, T_i) = \sigma_i$ that will induce 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

[7] 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 or development aid. The predictions of the model are thus also relevant at the level of federal government.

[8] This assumption simply says that rents that can be captured from still being in power next period are no bigger than the maximal rents that could be had today and that politicians discount future rents.

today to make abiding by the re-election rule more attractive than running away with maximum rents and forgoing re-election. This *fiscal restraint* constraint takes the form :

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

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

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

with T_i^* set such that the marginal value of the public good is equal to the marginal cost of taxation : $T_i^* = h(\frac{1-\alpha}{\phi})$. 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. Note that even with perfect information the fact that the incumbent can threaten to run away with all public revenues means he diverts rents in equilibrium.

2.2 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[9]. 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 to himself. A formal proof of this result is provided in the theoretical appendix, this section outlines the intuition.

Note first that to deter the incumbent in state H from implementing the L state menu the fiscal menus offered by the citizen must now also respect the incentive constraint :

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

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 the public good given the amount of taxes paid : state L's *fiscal restraint* constraint - equation (1)- is binding, so we have :

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

[9]This assumption is grounded in empirical evidence : Reinikka and Svensson (2005) for example show that local public funds coming from transfers are badly observed by citizens in Uganda and that improving information leads to less capture by local officials. What's new to this paper is the idea that taxes are better observed (here they are perfectly observed) because they come directly from the citizens' resources.

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 $W(G_H, G_L, T_H, T_L)$ subject to (5) and (6) determines the level of taxation in both states :

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

and

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

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 lower q) the more the citizen is willing to incur this cost, and the lower T_L^*. The theoretical appendix shows how it can be optimal for the citizen to pay no taxes in the L state for some values of the parameters. The asymmetry of information leads to an equilibrium with a higher level of taxation and 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 :

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

A marginal increase in taxes increases public good provision by $(1-\alpha)$ (assuming for simplicity that the increase does not affect the spread $T_H^* - T_L^*$) compared to $(1-\alpha)(1-u)$ for a marginal increase in average transfers, keeping everything else constant[10]. 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. Finally, note that the higher the asymmetry of information (higher u) the bigger the difference between 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 $\bar{f}^* = \bar{F}/E(R)$:

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

[10]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 low state.

I call this is the *accountability effect of taxes on the allocation of public spending* : 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, limits the extent to which a rent-seeking politician can capture public funds by 'hiding' them, and leads to an allocation of the budget that is more favorable to the citizen.

2.3 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[11]. This makes the 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 :

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

formal proofs of this expression and the ones that follow are found in the appendix. Because it decreases the share of transfers in total revenues f^* the reform also lowers the share of rents s^* (equation (10)). This leads to a first testable proposition regarding the impact of a tax administration reform:

Proposition 1 *A tax capacity program leads to an increase in taxes, an increase in public good provision and a decrease in the share of rents in total public revenues.*

A second proposition follows from comparing the impact of a rise in taxes of dT thanks to the program with that of an exogenous increase in the average value of transfer revenues($E(F)$) of the same amount :

Proposition 2 *The rise in taxes due to the reform leads to more increase in public good provision than a rise in transfer revenue of the same amount* $dE(T) = \frac{\partial E(T^*)}{\partial \phi} d\phi$ *:* $\frac{\partial G^*}{\partial E(T^*)} dE(T) > \frac{\partial G^*}{\partial E(F)} dE(T)$*. Reciprocally the rise in taxes will lead to a smaller increase in rents than an equivalent increase in transfers.*

[11] 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.

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 and the difference between taxes and transfers :

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 :* $\frac{\partial G^*}{\partial E(T^*)}dE(T) - \frac{\partial G^*}{\partial E(F)}dE(T)$ *is increasing in u.*

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

The first part of the empirical section of this paper offers a test of proposition 1 by evaluating the impact of a tax modernization program on public good provision and the share of rents (corruption) diverted by politicians in total public revenues. The second part tests propositions 2 by comparing the impact on public goods and corruption of an increase in taxes thanks to the tax modernization program to the impact of an exogenous increase in transfers and provides some evidence regarding proposition 3 by considering how these impacts vary when citizens have better access to information about the budget thanks to the presence of local media.

3 Context : Brazilian local governments and the PMAT tax capacity program

3.1 The Public Finances of Brazilian Local Governments and the PMAT program

The 1988 Brazilian constitution devolved substantial expenditure responsibility and tax autonomy to the country's more than 5000 local governments[12]. 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. This leads to a great diversity in tax revenues : in 2008 local governments collected anything from 2 to 2000 Reais per capita in taxes[13].

There is however a great difference between municipalities' *de jure* tax capacity and their *de facto* tax collection - they collect less than 13% of their total revenue themselves. This and the spiralling of local debts in the early 1990s has attracted much policy attention in Brazil and is often explained by the poor quality of local tax administrations. The few studies of Brazilian tax administrations available paint

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

[13]All statistics in this section are computed using the FINBRA database on local governments' public finances and program data from the BNDES and are in 2000 Reais. One Brazilian Real is equal to roughly 0.56 dollars.

a dire picture of unskilled and overworked staff with outdated tax registers, no institutional memory and a lack of methods to accurately assess tax liabilities. High costs of understanding and paying taxes for the citizens is thought to push many into non-compliance and local government officials have admitted to tolerating a situation of ongoing tax amnesty where tax arrears are rarely recovered[14].

The *Programa de Modernizacao de Administracao Tributaria* (PMAT) program was launched in 1998 by the Brazilian Development Bank (BNDES) to increase municipalities' capacity to tax their citizens. It provides local governments with subsidized loans to invest in modernizing their tax administration and is available to all municipalities in Brazil. Administrations prepare a detailed tax modernization project which is then assessed by the BNDES to check it qualifies with the program's requirement. Each local government was left free to choose the type of actions to take to modernize its tax administration but 'best practices' spread as the same firms or civil servants were involved in the development of several PMAT projects[15]. 330 municipalities started a program between 1998 and 2008, covering 40% of the Brazilian population. The paper's online appendix details the investments in tax administration financed by the program and several case studies.

Of particular interest to this study is the timing of application to and entry in the PMAT program. The amount of time between the date at which a municipality applies to a PMAT program and the start of a program lies between one and 4 years. This variation is due to the changing situation of the suppliers of the program. The BNDES processed all applications itself for the first 3 years of the program's existence. In order to reach more municipalities faster it decided in 2002 to contract the national bank *Banco do Brasil* to take in charge most of the application process[16]. *Banco do Brasil*'s involvement initially accelerated the application process until the bank decided to cut down the resources allocated to PMAT - it administered the start of 199 projects from 2002 to 2005, and 33 since. In 2007 the BNDES signed a similar agreement with another public bank, the *Caixa General*. 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 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 for a couple of years, and today the staff team has stabilized to around 12 individuals.

[14] 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)). See also Afonso and Araujo (2006) andBNDES (2002) for a discussion of local tax administrations in Brazil.

[15] For more on the program and the context of its creation see Santos et al. (2008), BNDES (2002) and Afonso and Serra (1999).

[16] The BNDES is based in Rio de Janeiro but the *Banco do Brasil* has branches around the country, allowing for more geographic outreach. The contract stipulates that *Banco do Brasil* help municipalities design a project that respects legal and financial rules and then transmit the application to the BNDES which is the only institution that can decide to start a program.

Table 1: Average Time Between Program Application and Start

Application Year	Years to Program Start	Nb Municipalities
1997	2.6	11
1998	2.7	10
1999	2.5	18
2000	2.3	13
2001	1.5	90
2002	1.1	29
2003	1	10
2004	0.7	33
2005	1.4	12
2006	1.8	11
2007	0.7	11
All	1.2	330

These evolutions are reflected in Table 1 which shows the average time between the year in which a municipality applied to the program and the year in which it started a program by year of application. It took 2 to 3 years to process an application for cohorts which started up to 2000 when the administrative resources devoted to PMAT were low. The process accelerated for the 2001 to 2004 cohorts, most of which were handled by the *Banco do Brasil*, and we see another decrease in process time for the 2007 cohort that applied in the year in which *Caixa General* was contracted. Whilst the possibility that the most eager municipalities pressured the BNDES to start the program soon after applying cannot be ruled out, in the overwhelming majority of cases (95%) the order in which municipalities applied to the program corresponds to the order in which they started a program. This suggests that whilst municipalities chose when to apply the precise date at which they started a program is mostly out of their control.

Local revenues which are not locally levied come from state and federal transfers. This study focuses on the transfer *Fundo de Participacao dos Municipios* (FPM) which is constitutionally mandated and the largest single source of local revenues (40%). This transfer has attracted the attention of researchers in the past because of nonlinearities in the rule allocating the distribution of *FPM* resources that provide exogenous variations which can be used to identify the impact of an increase in transfer revenues on outcomes of interest. Federal law defines 18 population brackets inside which the amount of *FPM* transfers a municipality receives is fixed, at each population threshold between brackets the amount received jumps by 20% on average. This *FPM* allocation rule has been used by two recent studies - Brollo et al. (2010) and Litschig (2008a). Using it is thus not a novelty of this paper, the interested reader is referred to these studies and the appendix for more details. Another advantage of considering *FPM* transfers for this study is the fact that its use is virtually unrestricted[17] : local

[17]Brazilian law requires that 15% of *FPM* revenues be spent on education and health. It is highly

politicians have the same discretion in deciding how to spend tax and *FPM* revenues. This implies that any observed difference in how marginal increases in these two revenues are spent cannot be due to any formal spending rule.

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. I use public finance data for the years 1999-2008 from the FINBRA dataset of the *Tesouro Nacional* to get the detailed revenue sources of local governments, in particular tax and *FPM* transfer revenues. All revenue variables used in the analysis are per capita, using official population estimates provided by the Brazilian statistical institute *IBGE*.

3.2 Expenditure responsibilities of Brazilian Local Governments

This project focuses on two aspects of the expenditure side of local governments' budgets : their spending on education inputs, and the extent to which public resources are wasted or diverted by local politicians. This is justified by data availability, the importance of both in the allocation of public revenue and their salience in Brazilian political debates.

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 in charge of primary schools (*ensino fundamental*)[18] Education is the largest budget item of local governments, representing nearly 30% of expenditures. I use panel data on municipal education inputs from the annual census (*Censos Escolar*) of Brazilian schools to measure the quality and quantity of municipal education infrastructure. The number of school classrooms in use in municipal schools (per thousand inhabitants) provides an indicator of how much local governments are effectively spending on education inputs. Classrooms are a component of school infrastructure which municipalities can easily adjust by furbishing existing unused rooms or renting extra office space. 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 of school transport and not on physical teaching infrastructure[19] which is therefore the education input most likely to be under-funded. 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, and with tv/video equipment.I use principal components analysis (PCA) to combine these four measures

unlikely that this rule ever binds however, given the importance of education and health as a share of total public budgets and the fact that there is no federal guidelines regarding what expenditure items can be labeled as health or education.

[18]By 2005, approximately 85% of all grade 1 to 4 schools were run by local governments (the remainder being private or state primary schools) who are responsible for providing school infrastructure, school lunches and transportation, and hiring, training, and paying teachers (see Ferraz et al. (2009)).

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

into an index of infrastructure quality. The first principal component explains 80% of the variation in the data, suggesting that using it reduces the dimensionality of the data with little loss of information.

The second public expenditure outcome this paper considers is corruption. There is considerable information on how local governments divert public resources away from public uses in Brazil thanks to an anti-corruption program launched by the federal government : 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 Uniao* (CGU) . These staff audit the use of discretionary federal transfers by local governments for the last two years 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 utilization of the two types of revenues -taxes and *FPM* transfers - is not directly audited, but because most discretionary federal transfers require that municipalities contribute some of their 'own revenues' (defined here as *FPM* transfers or taxes) on the programs they fund one can think of the audits as reflecting the overall quality of government spending. The results of those audits are publicly available records. Using those Ferraz and Finan (2009) 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[20].

Data on the extent of misappropriation/irregular use of public funds comes from the coding of the CGU audits for the years 2003-2006 provided by Litschig and Zamboni (2008). It is available for a small sample of 971 municipalities, 54 of which joined the program. Following the existing literature I construct a 'corruption' index from this data by scaling the number of irregularities by the number of potential 'offenders'- civil servants in the local government administration[21] - and the total amount of government revenue audited. This controls for the intensity of the search for irregularities undertaken by the CGU and provides a proxy for the share of diverted revenues in total government revenues, the theoretical outcome considered in the model above.

3.3 Why did municipalities join the program?

330 municipalities joined the PMAT program between 1998 and 2008 and selection was voluntary[22]. Interviews conducted with BNDES staff and local officials suggest the program was not of interest to the majority of Brazilian local governments whose small economic and population size make for a weak tax base and very low potential

[20]For more on the anti-corruption program and analysis using data from the audits see Ferraz and Finan (2008), Ferraz et al. (2009), Ferraz and Finan (2009),Litschig and Zamboni (2008) and Brollo et al. (2010).

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

[22]Program candidates were never directly rejected by the BNDES which only requires that funds are spend on modernizing the tax administration. Lack of proof of compliance to some federal regulations (for example the existence of overdue debt payments to a federal agency) did make some local governments ineligible.

returns to investments in tax administrations. Participants often said that they joined the program because they were dissatisfied with their current level of tax collection compared to what they thought was their tax potential. Low take up is also partially explained by the fact that the BNDES did very little promotion of its program : most participants said they heard about the program because they knew someone who worked at BNDES, or because one of the neighboring municipalities had already joined.

Given the large number of municipalities that did not join the program) I choose to exclude from my analysis those which fieldwork and inspection of summary statistics in Table 3 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, tax collection 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 35,000 (16% of municipalities which did not join the program), income per capita below 750 Rs (8%) and tax per capita in 1998 below 3.6 Rs (10%). Over 3,000 control municipalities remain to be included in the analysis.

To better understand why some local governments chose to join the program I estimate 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[23]. I look at the role played by pre-treatment values of GDP per capita, population (both estimated annually by the Brazilian statistical institute *IBGE*), and tax revenues and consider key demographic characteristics (median education, inequality and share of urban population, all from the 2000 Census) which could affect local political economy outcomes. The possibility that municipalities heard about the program from their participating neighbors is considered by including the average distance between a municipality and its 5 closest neighbors who have already joined a PMAT program and allowing for time dependence. Mayors with specific political ideologies and political ties may be more likely to join the program, so I include political party affiliation, alignment with the state governor's party and a measure of political competition to proxy for the type of political context mayors face[24]. I test the hypothesis that political or economic shocks influenced program uptake by including lagged changes in GDP per capita, tax revenues and whether a new mayor was elected during the previous election.

Results confirm the intuition gathered from field interviews : municipalities that join the program are rich and populated compared to the average Brazilian municipality but once these variables are controlled for they collect less taxes in 1998. They are also more educated, less agricultural and 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

[23]See Jenkins (1995) for a description of the method and Galiani et al. (2005) for a similar application to local water provision privatization in Argentina.

[24]All political variables are from the *Tribunal Superior Eleitoral*.

Table 2: Determinants of Program Uptake

	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 (d)	-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)
Growth in GDP		0.0303	
		(0.0873)	
Growth in population		-5.7851	
		(5.7434)	
Growth in taxes		-0.0000	
		(0.0000)	
Change in mayor (d)	-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

Hazard model of the probability of applying to the program : the dependent variable is a dummy equal to 0 for municipalities which have not applied to the program yet, 1 the year in which they apply, and missing hereafter. Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Municipalities which joined in the first two years are excluded because the variable 'distance to the 5 closest municipalities which have already joined a PMAT program' is not available for those.

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. Municipalities which have neighbors who have already joined the program are more likely to apply, possibly because they are more informed, and the coefficients for time dependance suggests information about the program spread gradually over time.

In the second column I test for whether past shocks determined program uptake, and find no evidence of an 'Ashenfelter dip' in tax revenues or that program participants were influenced by specific economic or demographic 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 neither economic nor political shocks determined entry in the program motivates the use of the difference in differences methodology described in the next section.

4 Impact of the tax modernization program on local tax revenues and spending outcomes

4.1 Empirical Strategy

I evaluate the impact of the tax modernization program on 1) local tax collection and 2) local expenditures outcomes that can potentially be financed from an increase in taxes, education inputs and corruption. In principle one would like to randomly assign treatment (program participation) to some municipalities and compare the average outcomes in the treated and control groups. Tax policies are hardly ever the subject of randomized trials[25] so I turn to non-experimental methods that create a credible counterfactual from the control municipalities under a reasonable set of assumptions.

The biggest identification concern is that treated and control municipalities could be different along dimensions which correlate with outcomes. For example richer municipalities join the program more often and they also collect more taxes and provide more public goods. Many of the unobservable characteristics that may confound identification are however likely to be fixed over time; I use panel data and estimate a difference in differences model to control for such time-invariant unobserved heterogeneity.

Formally, I estimate the model :

$$Y_{i,t} = \beta P_{i,t} + \delta X_{i,t} + \gamma_t + \mu_i + \epsilon_{i,t} \tag{13}$$

where $Y_{i,t}$ is either tax collected per capita in municipality i in year t or a measure of education inputs, $P_{i,t}$ is a dummy equal to 1 if municipality i is taking part in the program at year t and γ_t and μ_i a set of year and municipality fixed effect. I control for the key determinants of both tax collection and public services provision for which I have time-varying data, namely local population, GDP, political characteristics of the mayor (party, term limit) and the competitiveness of the last local election.

Panel data is not available for the corruption index so I estimate :

$$C_{i,t} = \beta P_{it} + \delta X_{i,t} + \delta_2 S_i + \delta_3 Z_i + \gamma_t + \epsilon_{i,t} \tag{14}$$

where $C_{i,t}$ is the corruption index, S_i is equal to 1 if municipality i joins the program between 1998 and 2008 and Z_i is a set of time invariant controls[26]. All specifications allow for arbitrary covariance structure within municipalities by computing standard

[25] An exception is Pomeranz (2010) which studies randomized enforcement of the VAT in Chile.

[26] I control for characteristics which are likely to affect corruption levels and/or program uptake: median education level, inequality and life expectancy from the 2000 census, whether the municipality is a state capital, whether tourism is a major industry, existence of a local radio stations and local judiciary presence.

errors clustered at the municipality level[27].

The key identifying assumption required for the interpretation of β in (13) and (14) as the average effect of the PMAT program is that the evolution of outcomes Y and C in treated municipalities would have been the same in the absence of the program as the evolution in 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 : if treated and control municipalities were not following similar trends before the program started it is unlikely that they would have in the absence of the program.

Table 2 already suggested 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. A more rigorous test that the pre-intervention time trends are the same is obtained by running a modified version of equation (13) using only observations for the pre-treatment period (1999-2008 for control municipalities, and years before joining for treated ones). Excluding the program dummy and including separate year fixed effects for (future) treated and controls I cannot statistically reject the hypothesis that pre-treatment year effects are the same for both groups[28]. This implies that tax revenues and public expenditure outcomes (education inputs and corruption) were following the same time trend in the pre-treatment period and is reassuring for the internal validity of my identification strategy. I present other tests which would detect violations of the identifying assumption in the robustness section below.

Because municipalities joined the program in different years I can also estimate a time-varying impact of the program :

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

where P_{jit} equals 1 if municipality i is in the jth year of the program in year t when $j \geq 0$ or will start the program in j years if $j < 0$. The β_j estimates are of interest for two reasons. First, they identify the dynamic impact of the program and estimate to what extent effects are sustainable over time. Second, estimates for $j < 0$ allow me for a test of an impact of the program 'before it happens'[29]. I run a test that the β_j

[27]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 allows for correlation amongst municipalities affected by similar shocks. Standard errors computed using this specification are slightly smaller suggesting spatial correlation is less of a concern than correlation over time within municipalities. Results using state-year clusters are available from the author upon request.

[28]Results available from the author upon request.

[29]Because of variation in the year in which the program started the fact that pre-treated year fixed effects are the same for treated and control groups could be overlooking pre-treatment trends that occur at different times for different cohorts.

are equal to zero for $j < 0$ to complement the test on year fixed effects explained above.

A final concern is that pretreatment characteristics that are thought to be associated 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 treated municipalities and controls. 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)). In this case simple 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 source of bias 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 (13) 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 of common support in the distribution of covariates, and 2) obtaining balance of covariates by re-weighting the control group observations. 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 2, obtaining the predicted probability $\hat{P}(W)$ and then estimating (13) with weights equal to unity for the treated and $\hat{P}(W)/(1-\hat{P}(W))$ for the controls.Hirano et al. (2003) show that this estimator is efficient. More detail on the model used to estimate the probability, the distribution of the propensity score and the selection of the common support sample are in the appendix[30]

4.2 Summary statistics and graphical evidence

Table 3 presents summary statistics of key characteristics of treated and control municipalities. Treated municipalities levied 95 Reais per capita in 1998, nearly three times more than control municipalities. They were also much bigger (180,000 inhabitants on average) and consequently received less *FPM* transfers per capita. Average municipal revenues per capita were however very similar across the two groups. Municipalities that join the program had in 1998 roughly the same amount of school infrastructure but better school quality, and the randomized audits reveal a lot less irregularities for them than in control municipalities on average over the period. Other municipality characteristics are as expected given the results for the selection model presented in table 2. The last column of the table shows that weighting control observations by a function of their propensity score obtains a reasonable balance of pre-treatment char-

[30] **?** notes that implementing this method whilst ignoring the impact of the first-stage estimation of the selection probabilities on standard errors obtains conservative inference. All results below therefore present standard errors non-adjusted for first stage estimation. The use of a bootstrap method yields very similar standard errors.

acteristics across the treated and control groups.

The evolution of tax revenues in control and treated municipalities from 1998 to 2008 is presented in Figure 1 in which tax revenues in treated municipalities are scaled so that the year in which the program starts is assigned to 2002. We see a clear increase in tax collection in the treated municipalities once the program has started, however the most striking element in figure 1 is probably the great difference in tax collection between the treated group and the group composed of all control municipalities. This difference in levels makes it difficult to assess the validity of the common trend assumption. The evolution of the common support, or 'small' control group, constructed as explained above and in which the observations are weighted according to their propensity to join the program, appears to provide a more credible counterfactual. Tax collection levels are very similar for this group and the treated municipalities in 1998 and both groups follow similar trends until 2001. Comparing their evolution post 2002 suggests the program increased tax collection by 10-15% after 3 years. They diverge increasingly over time but part of this divergence may be due to a sample composition effect, as only municipalities which started the program early are used to compute average tax collection towards the end of the period. Regression analysis gets rid of this composition effect, and controls for time-varying covariates that differ across treated and control municipalities.

4.3 Results : Impact of the program

Table 4 reports results from the estimation of equations (13) and (15) using tax revenues per capita as the dependent variable. The first column presents results for a model using the whole sample and including no covariates except for municipality fixed effects and year dummies. I find that the program is associated with a 10.5 Rs increase in tax revenues, which amounts to a 11% increase in tax collection compared to the baseline 1998 level in treated municipalities. Controlling for time-varying covariates in column 2 and restricting the sample to the common support in column 3 decreases the estimate slightly. Finally the estimated impact of the program is an extra 7.7 Rs of taxes per capita (a 8% increase) when using the preferred propensity score weighted regression method. This suggests that the difference in pre-treatment characteristics between treated and control municipalities leads to some over-estimation of the impact of the program. Results in all the following tables are therefore obtained using the propensity score weighted method, estimates from alternative specifications are presented in the appendix.

Columns 5 6 and 7 show results for the dynamic impact of the program using the weighted difference in differences model in equation (15). I restrict for now the β_j to be equal to 0 for $j < 0$ to obtain more efficient estimates. I only report results estimated on a fixed number of municipalities, to avoid confounding the estimation of the program's dynamic impact with that of potential heterogenous effects depending on time of entry in the program[31]. The sample used to provide the estimates in column

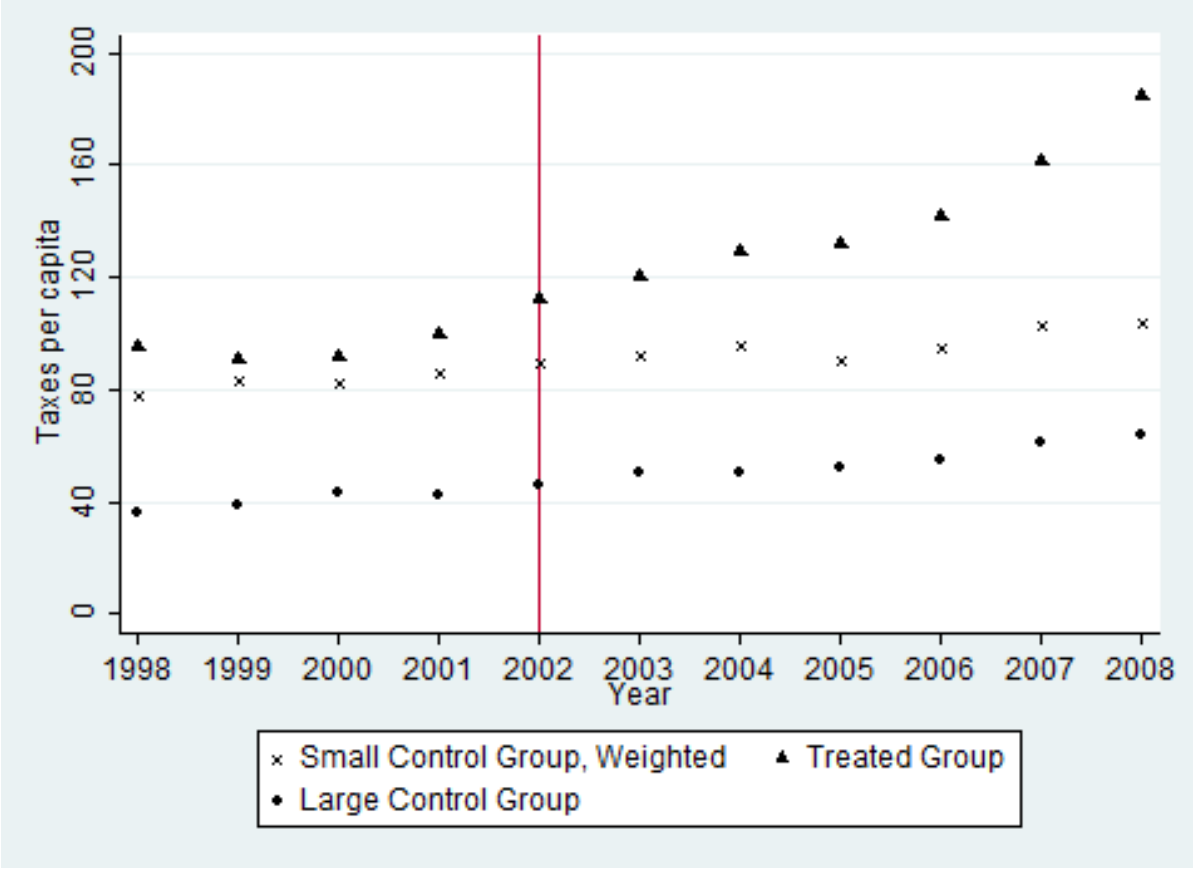
[31]The specifications includes a separate dummy for each year in the program.

Table 3: Summary Statistics

		Unweighted sample		Weighted sample	
	Treated Municipalities	**Control Municipalities**	**All municipalities**	**Control Municipalities**	**All municipalities**
Pre-treatment values					
Taxes (1998)	95.17 (97.49)	38.92 (79.46)	44.11(82.89)	78.40(66.65)	81.73(74.08)
FPM Transfers (1998)	93.9 (63.9)	167.2(76.1)	160 (78.04)	120.1 (56.5)	116.9(62.6)
Total revenues (1998)	449 (193)	440 (208)	441 (222)	565 (192)	538(197)
School infrastructure (1998)	4.64 (2.54)	4.18 (2.13)	4.04 (2.13)	5.01(2.24)	4.54(2.31)
School quality (1998)	0.4 (1.2)	-0.28 (0.4)	-0.22 (1.6)	0.2(1.5)	0.3(1.4)
Population (1998, thousands)	180 (70)	24 (57)	40 (23)	150 (46)	160 (52)
GDP (1999)	7.09 (5.14)	4.45 (4.12)	4.71 (4.30)	8.23(3.77)	8.02(4.09)
Agr\ GDP (1999)	10.15 (12.24)	23.74(15.75)	22.43 (15.96)	36.94(23.54)	31.86(24.24)
Serv\ GDP (1999)	63.39 (12.61)	59.48 (14.13)	59.85 (14.04)	45.81(21.12)	49.15(20.95)
Income	2.34 (1.16)	2.21 (1.05)	3.56 (1.48)	3.20 (1.09)	3.27(1.19)
Census variables (2000)					
Inequality	0.55 (0.05)	0.55(0.05)	0.55 (0.05)	0.56(0.03)	0.56(0.04)
Urban pop.	0.84(0.12)	0.64 (0.22)	0.66 (0.22)	0.8(0.11)	0.80(0.12)
Education	5.65 (1.27)	4.25 (1.17)	4.38 (1.25)	5.07 (1.13)	5.18(1.19)
Average for whole period					
Corruption	74.50 (171.86)	115.31(159.23)	112.05(160.53)	46.13(84.46)	59.13(129.05)
Has at least one AM radio in 1998	0.67 (0.47)	0.23(0.42)	0.27(0.45)	0.23 (0.42)	0.31(0.46)
Change in mayor	0.14(0.35)	0.15(0.35)	0.15(0.36)	0.11(0.32)	0.12(0.33)
Governor's party	0.23(0.42)	0.22(0.41)	0.23(0.42)	0.11(0.31)	0.13(0.34)
Pol. competition	0.57(0.10)	0.52(0.11)	0.53(0.12)	0.58(0.11)	0.57 (0.11)
Distance to closest PMAT	311.70 (449.48)	356.08 (413.82)	351.95(417.46)	291.54(372.62)	295.93(387.18)

Columns 5 and 6 restrict the sample to the common support sample and weight observations according to their propensity score. All revenue variables are per capita. School infrastructure is the number of classrooms in use in municipal schools per 1000 inhabitants, school quality is an index constructed from several indicators of school quality using principal component analysis, the political competition variable is the Herfindahl index of the concentration of votes during the last municipal elections, and distance to closest PMAT is the average distance between the municipality and the 5 closest municipalities who have already joined a PMAT program.

Figure 1: Evolution of Tax Revenues in Treated and Control Municipalities



5 therefore excludes municipalities who started the program after 2005 because those are observed less than 4 years after they join, in column 6 I exclude municipalities who started after 2004 , and in column 7 those who started after 2003. The immediate impact of the program is small (5% increase) but it reaches 10-11 Rs per capita after 4 years. This 11% increase in tax collection seems sustainable : estimates vary little from the 4th year onwards. On average municipalities borrowed 9 Rs per capita at a real interest rate of 5% through the program and reimbursed the loan after 6 years, leading to a total cost of 12 Rs per capita. The program was highly cost-effective: on average a municipality recovered invested funds after 2 years in the program; after 4 years one Real invested in modernizing the tax administration leads to slightly more than one extra Real in tax revenue in each year.

Table 4: Impact of the Program on Tax Revenues

	1	2	3	4	5	6	7
	DiD	DiD	DiD on common support		Weighted DiD		
All years	10.523***	9.946***	9.204***	7.243***			
	(1.945)	(2.093)	(2.600)	(2.589)			
1th year					5.170*	4.978*	4.783
					(2.671)	(2.744)	(3.021)
2th year					8.148**	8.321**	6.443
					(3.488)	(3.677)	(3.923)
3th year					10.065***	9.907***	9.407**
					(3.264)	(3.431)	(4.299)
4th year					11.164***	11.628***	11.063**
					(3.330)	(3.435)	(4.427)
5th year						10.500***	10.589**
						(3.835)	(4.533)
6th year							11.435**
							(5.400)
Observations	35600	35562	24049	24049	23416	23290	22300
Municipalities	3654	3654	2462	2462	2397	2384	2283

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Tax revenues are measured per capita. All regressions include municipality and year fixed effects, controls in columns 2 to 5 are 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 of the impact of the program on local expenditure outcomes are presented in Table 5. The table shows that the program leads to the opening of an extra 0.2 classrooms per 1000 inhabitants on average. The dynamic impact of the program mirrors that of the increase in tax revenues described above, with an extra 0.4 classrooms after 5 years in the program, an 8% increase with respect to the baseline. The school quality index increases by around 30% compared to its baseline value, which

corresponds to one-tenth of a standard deviation amongst the treated group. Finally column 6 suggests that the corruption index which proxies for the share of resources diverted in total public revenues has decreased thanks to the program. The estimate is very large (it corresponds to a 50% decrease) but imprecisely estimated.

It is not a priori clear how to interpret Table 5 as the increase in the quantity and quality of municipal education supply may have been financed by either the extra tax revenues generated by the program, or by the loosening of the budget constraint allowed by the program loan itself : if treated municipalities would have invested in their tax administration without the program, the program money could have allowed them to allocate extra funds to education expenditures. A rough back of the envelope calculation suggests the latter explanation does not hold. Both treated and control municipalities had local public revenues of around 450 Rs per capita in 1998, and financed the use of around 4.6 classrooms per 1000 inhabitants. Assuming that the average and marginal propensity to spend on opening new classrooms out of an increase in municipal revenue are equal, one would expect municipalities to have to obtain an extra 100 Rs per capita before they open one new classroom (per thousand inhabitant). The average loan amount, 9 Rs per capita, clearly would not suffice to open up to 0.4 extra classrooms unless we're willing to assume extremely large differences between the marginal and the average willingness to spend on classrooms. On the other hand the program's cumulated impact over 5 years is to increase local revenues by 50 Rs. The idea that this amount was enough to open 0.4 extra classrooms seems reasonable. The same logic applies to investments in school quality. Overall Tables 4 and 5 indicate that the program increased the amount of tax revenues available to treated municipalities and also increased the availability of the main type of public good financed with local revenues - education - without leading to more corruption.

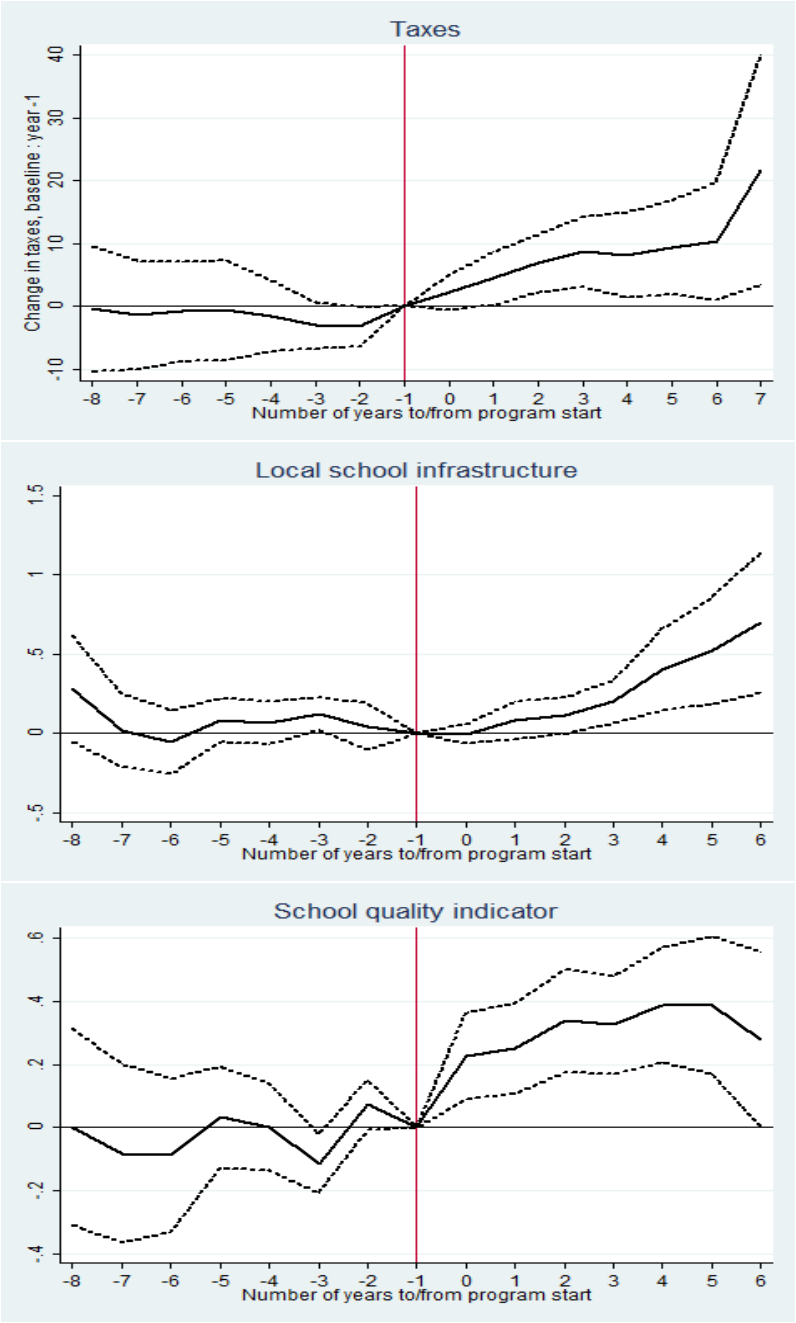
Finally Figure 2 provides some reassuring evidence regarding the validity of the common trend assumption necessary for identification by graphing the unrestricted dynamic estimates obtained from estimating equation 15. Each point on the solid lines summarizes the effect of having been in the program 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 fiture shows that these effects are never statistically different from zero before the program starts. This confirms that treated municipalities did not join the program at a time when they had already started to increase tax revenues and public goods.

Table 5: Impact of the Program on Local Expenditure Outcomes

	Infrastructure				School quality				Corruption
All years	0.206***				0.142**				-35.232*
	(0.069)				(0.059)				(20.546)
1th year		0.047	0.040	-0.095		0.081	0.085	0.125	
		(0.040)	(0.043)	(0.093)		(0.058)	(0.063)	(0.078)	
2dyear		0.101*	0.109*	0.181***		0.117**	0.124**	0.117**	
		(0.058)	(0.064)	(0.089)		(0.058)	(0.061)	(0.087)	
3d year		0.152	0.127	0.159**		0.158**	0.169**	0.188***	
		(0.091)	(0.077)	(0.075)		(0.065)	(0.067)	(0.096)	
4th year		0.202***	0.160**	0.230**		0.173**	0.186**	0.186***	
		(0.072)	(0.072)	(0.111)		(0.073)	(0.073)	(0.096)	
5th year		0.311***	0.405***	0.270***		0.189*	0.192*	0.264***	
		(0.134)	(0.134)	(0.099)		(0.098)	(0.099)	(0.104)	
6th year		0.425***	0.419***	0.391***		0.240*	0.243*	0.301**	
		(0.174)	(0.173)	(0.153)		(0.130)	(0.131)	(0.130)	
Observations	24053	23419	23292	22302	24053	23419	23292	22302	483
Municipalities	2462	2397	2384	2283	2462	2397	2384	2283	483

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Infrastructure is number of classrooms in use in municipal schools per 1000 inhabitants, school quality is the first principal component constructed from the school quality variables described above, and corruption is the misgovernance index compiled from the CGU audits and Litschig and Zamboni (2008).All regressions include municipality and 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. The specification used in the last column also controls for state fixed effects, lottery fixed effects, median education level, inequality and life expectancy, whether the municipality is a state capital, whether tourism is a major industry, existence of a local radio station and local judiciary presence.

Figure 2: Year by Year Impact of the Program



4.4 Robustness checks

The above analysis uses the evolution of outcomes in municipalities that have not joined the program yet, including those who are never observed starting the program, as a counterfactual outcome to identify the impact of the program on the treated municipality. An alternative is to use only the evolution of outcomes in municipalities who have not joined yet at time t, but will later at a time $t+s$ observed in the data, as those are arguably very similar to the municipalities who have already joined at time t. Unfortunately the bunching of municipalities' program start date around a few main years makes it impossible to estimate equation (13) on a sample consisting only of the 330 municipalities who enter the program before 2009 and identify separately year fixed effects and the program's impact (see appendix Table 5). Table 6 nevertheless presents estimates obtained using this sample which approximate what one would ideally like to do. In the first column I present estimates obtained from estimating equation (13) using the sample of treated municipalities only and up to the year 2003 : in this sample the 187 municipalities which start the program after 2003 are never in the program, and are therefore used as control municipalities for the 147 who are treated up to 2003. The estimated impact of the program on this smaller sample is close to what we obtained above, though less precisely well estimated. The second column shows results from estimating equation (13) on the sample of treated municipalities only and using all the years available but without year fixed effects. Whilst the estimated 'program impact' confounds the true program impact and the increase in outcomes over time that is common to all municipalities the comparison of these estimates with those obtained from estimating the same specification (equation (13) without year fixed effects) on the main sample used for analysis in column 3 is of interest. The estimates in columns 2 and 3 could differ for two reasons which would be a concern for the identification strategy used above; first if the evolution of outcomes before the program starts was different in control and treated municipalities and second, if I mis-specified the impact of time-varying covariates that are still unbalanced between treated and control municipalities despite the re-weighting. We have already seen evidence that the first cause for concern does not hold in the data, but the fact that the estimates in columns 2 and 3 are remarkably similar provides further reassurance regarding the validity of the identification strategy used.

The remaining concern regarding the interpretation of the estimates as impact of the program is that of an unobservable shock occurring at the same time as the program and affecting local outcomes. For example the characteristic 'government capability and honesty' could increase for unobserved reasons in one year, leading local officials to simultaneously apply for the program, become more efficient at collecting taxes and more accountable in their allocation of the budget. Whilst the assumption that this does not happen is essential to my identification strategy and cannot be tested I use the arguably exogenous time lag between program application and program start to offer some evidence that such unobservable shocks do not explain the observed changes in outcomes after the start of the program.

As described above the precise date at which a municipality joins the program is

Table 6: Impact of the Program, Alternative Specifications

	Treated only, until 2003	Treated only, no year FE	Whole sample, no year FE
Impact on taxes			
Program : all years	7.746*	16.802***	16.846***
	(4.630)	(1.747)	(2.509)
Observations	959	3188	24049
Municipalities	330	325	2462
Impact on school infrastructure			
Program : all years	0.161***	0.208***	0.204***
	(0.051)	(0.037)	(0.051)
Impact on school quality			
Program : all years	0.113*	0.385***	0.367***
	(0.064)	(0.052)	(0.061)
Observations	960	3195	24053
Municipalities	330	325	2462

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Tax revenues are per capita, infrastructure is the number of classrooms in use in municipal schools per 1000 inhabitants and school quality is the first principal component constructed from the school quality variables described above. All regressions include municipality 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; year fixed effects are used in column 1.

partially out of its control. If municipalities apply to the program because their local officials become more capable and honest and this has an impact on outcomes we should see a change in outcomes at the time a municipality applies *even if the program itself does not start for a couple of years*. Table 5 presents estimates of the average impact of the program and the variation in local outcomes for the 3 years prior to program start for municipalities which apply and start a program in the same year (63 municipalities), those that wait one year (174) and those that wait 2 or 3 years (81).

We see that for all these sub-groups and all outcome variables the variation in outcomes prior to program start is not at all statistically significant. It is in particular impossible to see a jump in outcomes at the date of application when municipalities apply 2 to 3 years before joining (third column). Of course we expect to see no such pattern if the municipalities whose files were processed slowly are also the ones who were the less eager to join the program and hence experienced a smaller and more delayed impact of the program. The average impact of the program is similar for all groups however; this suggests that any heterogeneity in the treatment effect is not correlated with the speed of application.

Finally Appendix table 13 presents results for the dynamic impact of the program using different specifications and sample size. When the whole control group is used and observations are not weighted according to their propensity score the program seems to have a larger impact on taxes, municipal school infrastructure and quality. Though estimates remain close to the ones obtained using the preferred specification this confirms that restricting the sample to construct a more likely counterfactual

Table 7: Impact of the Program by Lag between Application and Program Start

	0	1	2-3
Impact on taxes			
3 years before	3.254	-3.985	-0.524
	(4.855)	(3.078)	(3.081)
2 years before	0.466	-4.792	-0.405
	(5.728)	(3.585)	(3.358)
1 year before	-0.035	0.276	0.467
	(5.947)	(4.425)	(3.958)
Program : all years	7.979*	6.794*	8.114*
	(4.129)	(3.388)	(4.674)
Impact on school infrastructure			
3 years before	0.219	-0.070	0.130
	(0.159)	(0.064)	(0.115)
2 years before	-0.041	-0.138	-0.004
	(0.165)	(0.096)	(0.129)
1 year before	-0.002	0.026	-0.061
	(0.187)	(0.185)	(0.131)
Program : all years	0.290*	0.155*	0.248*
	(0.172)	(0.092)	(0.128)
Impact on school quality			
3 years before	0.031	-0.017	0.009
	(0.082)	(0.073)	(0.110)
2 years before	0.060	0.106	0.055
	(0.095)	(0.111)	(0.122)
1 year before	0.127	0.042	0.071
	(0.108)	(0.081)	(0.114)
Program : all years	0.190*	0.116*	0.213**
	(0.110)	(0.069)	(0.074)
Observations	21619	22642	21710
Clusters	2214	2318	2224

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Tax revenues are per capita, infrastructure is the number of classrooms in use in municipal schools per 1000 inhabitants and school quality is the first principal component constructed from the school quality variables described above. 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.

matters, as suggested by figure 1. A recurring criticism of the difference-in-differences methodology is that it is strongly functional form dependent (Heckman (1996)). I present results for the natural logarithm of the dependent variables using the preferred propensity-score weighted method which paint a similar picture of the dynamic impact of the program[32].

The availability of rich panel data and the exploitation of the program's timing thus allows me to rule out the most likely alternative hypothesis that could explain the observed changes in outcomes once the program has started and argue that the program itself has a causal impact on tax revenues, which are then used to finance extra education provision but no increase in corruption. This confidence with respect to causality however does not imply that the program would lead to such outcomes if applied to all Brazilian municipalities. Two things are required to increase tax collection : the program's money and technical knowledge, and motivation of local politicians and administrations. The last set of results suggest it is unlikely that local motivation (proxied by the timing of application to the program) is a sufficient condition for the observed change in outcomes, however it is likely to be a necessary condition. Imposing the program on municipalities in which local officials are not interested in increasing tax collections probably would not work, in that sense all the above estimates must be understood as impact of the program *on the treated*, on the municipalities which had self selected into treatment. To provide estimates of the impact of the program were it imposed on the average Brazilian municipality one would need to assume knowledge of the local motivation to increase in tax capacity in this municipality, an assumption that I am not willing to make. This interpretation of the estimates as 'treatment on the treated' must be kept in mind when comparing the marginal impact of an increase in taxation to that of an increase in transfers in the next section.

[32]Results differ when using a log specification and the large sample. This is a consequence of the large difference in levels between the treated and the control groups when the latter is not restricted to the group on which there is common support, which motivates the combination of a difference in differences specification and propensity score matching.

5 Are tax revenues better spent than transfer revenues?

This section compares the impact on local expenditure outcomes of an increase in tax revenues thanks to the program to the impact of an exogenous increase in transfer revenues. It directly tests proposition 2 of the model which states that an increase in taxes thanks to the program will lead to a higher increase in public good provision, and a smaller increase in corruption, than an increase in transfers of the same amount.

As explained above the impact of an exogenous increase in transfers on corruption and education provision at the local government level in Brazil has already been studied in the literature by studies that exploit discontinuities in the allocation rule of the largest unrestricted transfer to local governments, the *FPM* transfer. Exploiting these discontinuities in a regression discontinuity design framework Brollo et al. (2010) find that an exogenous increase in transfers of 10% raises incidence of corruption by 12 percentage points. Litschig (2008a) estimates that a 1000 Rs increase leads to 0.42 more years of elementary schooling and a 5.6% increase in student literacy rate.

In contrast the results in the previous section suggest, using a simple Wald estimate, that a 10% increase in taxes thanks to the program leads to a (statistically insignificant) fall in the share of resources diverted. I cannot compare my findings with those in Litschig (2008a) because his data on education outcomes is not available for my period of study. I therefore replicate both studies' estimation strategy in attempt to directly compare the impact of tax and transfer revenues on corruption and education inputs.

5.1 Empirical Strategy

I evaluate the impact of increase in taxes thanks to the program by estimating the following equations :

$$E_{i,t} = \pi_{TE} T_{it} + \eta X_{i,t} + \gamma_t + \mu_i + \epsilon_{i,t} \tag{16}$$

and

$$C_{i,t} = \pi_{TC} T_{it} + \eta X_{i,t} + \eta_2 Z_i + \eta_3 S_i + \gamma_t + \epsilon_{i,t} \tag{17}$$

where $E_{i,t}$ is a measure of education inputs and $C_{i,t}$ the corruption index, $T_{i,t}$ is instrumented for using program participation and all covariates are as above. The identifying assumptions required to make a causal interpretation of the IV parameters π valid are the same as those used to interpret the difference-in-differences estimates : there must be no unobserved time-varying municipal characteristics that affect program uptake, tax revenues and local expenditure outcomes simultaneously. The discussion above of those assumptions is therefore still relevant in this section.

The impact of intergovernmental transfers is identified using the exogenous variations in the amount of FPM grants received generated by the transfer allocation rule outlined in the appendix. This rule specifies that all municipalities in the same state and in a given population bracket should receive the same amount of transfers.

Appendix Figure CHECK 3 shows that, although there are multiple cases of misassignments around the population thresholds, the amount of FPM transfers received by municipal governments displays clear jumps at each threshold. Following Brollo et al. (2010) I therefore use a fuzzy regression discontinuity approach, using the amounts that the rule predicts each municipality should receive (theoretical transfers) as an instrument for the grants actually received and controlling for a high-order polynomial in population size. The estimated equation for transfers is :

$$Y_{i,t} = \pi_{FY} F_{i,t} + \eta_{FY} X_{i,t} + \gamma_t + \mu_i + v_{i,t}, \tag{18}$$

where $Y_{i,t}$ is education inputs or the corruption index[33], $F_{i,t}$ is FPM resources per capita, instrumented for using theoretical FPM per capita and I flexibly control for population size by allowing for different slope and curvature around the thresholds. The identification of the impact of transfers in equation (18) comes from both cross-sectional variations (municipalities just below and just above the thresholds) and within-municipalities variations (municipalities who cross the thresholds over time). Because the variation used to identify the impact of taxes in equation (16) is within municipalities I present estimates of the transfer equation with and without municipality fixed effects and show that results are similar. More details on the methodology used is in the Appendix SAY MORE.

Finally equations 17 and 18 using the corruption index as the dependant variable are estimated using the two-sample instrumental variable (TSIV) method developped by Angrist and Kruger (1992, 1995) (see also Inoue and Solon (2010)). TSIV is appropriate in situations in which the outcomes are available in one data set, the endogenous regressor is available in a second data set, both data sets contain the instrumental variable and the other exogenous variables included in the model, and the distribution of variables is the same in both datasets. In my case the indicator for participation in the PMAT program and tax variable are both present in the large sample of nearly 3000 municipalities, but the corruption index is only available for a random subsample. The TSIV method increases the precision of the first stage by first estimating the impact of the PMAT program and theoretical transfers on tax and transfer revenues on the large sample, then predicting tax and transfer using the first stage estimates, and finally using these predicted variables as regressors in the second stage equation on the small sample for which the corruption index is available. Standard errors taking into account the fact that these second stage regressors are estimated are obtained using bootstrap.

5.2 Results

Table 8 presents estimates of the marginal impact of taxes and transfers on local education infrastructure and school enrollment in the top two panels. F statistics for the corresponding first stage estimates are reported to assess the strength of the (conditional) correlation between the instruments and the endogenous variables. The first

[33]When $Y_{i,t}$ is the corruption index municipality level fixed effect is replaced by a treated fixed effect and the set of time-invariant covariates used above.

column reports results from estimating equations (16) : estimates are as expected very close to what we obtain by computing Wald estimates from Tables 4 and 5. The second and third column show the marginal impact of a rise in transfers estimated with the model in equation (18), with or without year fixed effects and suggests that increases in transfer revenues has little, if any, impact on municipal education supply. Though the identification strategy used for the impact of transfers seems more cross-sectional in spirit the F test for the first stage of the IV estimates using only variations over time is still very high (48) : the transfer allocation rule appears to lead to substantial variations in transfers even within municipalities over time.The last column therefore estimate the impact of taxes and transfers jointly using the preferred fixed-effect and propensity score weighted specification (column 5). Increases in taxes thanks to the program lead to more investment in school infrastructure and a higher increase in school enrollment than the increase in transfer revenues that occurs thanks to the *FPM* allocation rule, in line with the model's proposition 2. The increase is anywhere between 2 and 20 times bigger for taxes than for transfers.

The last panel of Table 8 presents results regarding the impact of a rise in taxes or transfers on the corruption index. Compared to the (weighted) baseline value of 56 for the corruption index we see that a 1 Rs increase in transfers leads to a nearly 3% increase in corruption, an estimate that is very close to those obtained in Brollo et al. (2010). The last column suggests that an increase in tax revenue has the opposite impact, in line with the model's predictions.

The extent to which the estimates of the impact of extra transfers or taxes on local expenditure outcomes are comparable needs to be discussed. An ideal test of the model's prediction 2 that increase in taxes lead to a more favorable allocation of public revenues than increases in transfers would require the existence of two perfectly identical municipalities A and B, one in which local taxes are exogenously increased, and one in which transfer revenues are exogenously increased by the same amount. The research design used in this paper to attempt to compare the impact of taxes and transfers necessarily differs from this ideal design. As explained above it is likely that municipalities which enter the program, and so are affected by the tax instrument, are different from the average Brazilian municipality along an unobservable dimension which could be correlated to its government's propensity to spend on education and/or be corrupt. Whilst the discussion in section 4 suggests that the estimated coefficients in Table 8 can be interpreted as the causal impact of an increase in taxes on local expenditure outcomes, these coefficients are necessarily *local* causal impacts. They represent the impact of an increase in taxes thanks to program participation in the municipalities that are willing to join the program. Nothing ensures that these municipalities are similar to the ones affected by the instrument for transfer revenues, the population discontinuity threshold, and so that the comparison between the estimated coefficients really captures the theoretical mechanism developed in the model.

Weighting control observations by their propensity score means that the impact of an increase in transfer revenues is estimated on an average municipality that is similar to the treated municipality along observable characteristics. The comparison between

Table 8: Impact of Increases in Tax or Transfer Revenues on Local Expenditure Outcomes : IV Estimates

	Fixed Effects	No Fixed Effects	Fixed Effects	Fixed Effects
Impact on school infrastructure				
Taxes	0.026***			0.028**
	(0.010)			(0.013)
Transfers		0.011	0.007	0.005
		(0.007)	(0.007)	(0.006)
Impact on school quality				
Taxes	0.016**	0.015***		0.015***
	(0.007)	(0.005)		(0.005)
Transfers		0.001	0.003	-0.001
		(0.002)	(0.001)	(0.002)
F stat (Taxes)	12.8			13.2
F stat (Transfers)		441	48.2	50.1
Observations	24049	22673	22675	22673
Municipalities	2462	2331	2333	2331
	No Fixed Effects	No Fixed Effects	No Fixed Effects	
Impact on corruption				
Taxes	-4.267		-4.471	
	(2.709)		(3.502)	
Transfers		1.633***	1.604***	
		(0.404)	(0.425)	
F stat (Taxes)	10.8		11.5	
F stat (Transfers)		441	450	
Observations	483	462	462	
Municipalities	483	462	462	

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.Infrastructure is number of classrooms in use in municipal schools per 1000 inhabitants, school quality is the first principal component constructed from the school quality variables described above and corruption is the misgovernance index compiled from the CGU audits and obtained from Litschig and Zamboni (2008). 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. The specifications in which municipality fixed effects are not included also control for state fixed effects, median education level, inequality and life expectancy, whether the municipality is a state capital, whether tourism is a major industry, existence of a local radio station and local judiciary presence. When the corruption index is the dependent variable the estimates are obtained using TSIV, and standard errors are estimated using bootstrap.

the estimates for tax and transfer revenue increase in Table 8 is therefore only as credible as the assumption that this average municipality is comparable to the treated municipalities along unobservable characteristics as well. Note however that these estimates do provide a rigorous answer to one policy question of interest : they suggest that if the Brazilian federal government is interested in maximizing the amount of local education provided by municipal governments on average and minimizing their diversion of public revenues providing municipalities with subsidized loans to invest in their tax capacity is much more cost-effective than increasing the constitutional transfers that every local government in Brazil is entitled to.

Table 9 offers a test of the model's prediction that the marginal impact of tax and transfer revenues on public good provision and corruption should be more similar the more information citizens have about the public budget (proposition 3). 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. The model predicts that the difference between the impact of a marginal increase in taxes and transfers should be smaller in municipalities with a local radio stations; table 9 reports the p-values of t tests for the hypothesis that the impact of taxes is equal to the impact of transfers in municipalities with and without a radio station. The evidence regarding the quantity and quality of local education supply (columns 1 and 2) suggests that information does make the accountability inducing advantage of tax revenues relative to transfer revenues smaller, as the difference between the two sources of revenue becomes statistically insignificant. When we use the corruption index as a proxy for the share of rents in total revenues the evidence is inconclusive however : once terms interacting revenue increases with presence of local radio are introduced all estimated coefficients loose statistical significance, and one cannot distinguish different impacts of taxes and transfers on corruption with or without the presence of a local information medium.

Appendix Tables 14 and 15 explore two alternative mechanisms which could lead to a difference in how increases in taxes and transfer revenues are spent. I test for the possibility that governments which rely more on local tax revenues have better incentives to invest in public goods if they increase the local tax base by considering the program's impact on local GDP and population. Results show that neither variable is affected by program participation, suggesting this mechanism is not relevant in the context of Brazilian local governments, possibly because the types of investments local governments can make in Brazil (in education or health) are unlikely to affect local growth fast enough to be a relevant factor for politicians and be detected in the data. Another difference between taxes and transfers could be that tax revenues are more stable than transfer revenues. If local governments only invest in education quantity and quality when they experience an increase in revenues that they believe is stable over time, and divert increases in revenues that are short lived, this difference could explain the results in Table 8. The within municipality standard deviation is always smaller relative to the mean for transfer revenues than for tax revenues, suggesting that this mechanism does not apply either. This is unsurprising given that the transfers considered - *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).

This mechanism may however be relevant if one considered discretionary transfers, which are more volatile.

Table 9: Impact of Taxes and Transfers with and without local media

	Infrastructure	Quality	Corruption
Taxes	0.040*	0.014*	-3.060
	(0.024)	(0.008)	(32.967)
Transfers	0.004	-0.001	7.681
	(0.007)	(0.002)	(19.459)
Taxes*radio	-0.024	0.001	-17.219
	(0.020)	(0.007)	(39.199)
Transfers*radio	0.013	0.009	-5.851
	(0.010)	(0.008)	(75.583)
T-test p value (no radio)	0.06	0.02	0.8
T-test p value (with radio)	0.91	0.45	0.9
Observations	22673	22673	462
Municipalities	2331	2331	462

Cluster-robust standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Infrastructure is number of classrooms in use in municipal schools per 1000 inhabitants, school quality is the first principal component constructed from the school quality variables described above and corruption is the misgovernance index compiled from the CGU audits and obtained from Litschig and Zamboni (2008). 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. The specifications in which municipality fixed effects are not included also control for state fixed effects, median education level, inequality and life expectancy, whether the municipality is a state capital, whether tourism is a major industry, existence of a local radio station and local judiciary presence. When the corruption index is the dependent variable the estimates are obtained using TSIV, and standard errors are estimated using bootstrap. The variable 'radio' is an indicator equal to 1 if the municipality has at least one local radio station.

6 Conclusion

This paper provide optimistic evidence on the impact of a tax modernization program on Brazilian local governments - a permanent 11% increase in taxes per capita, more local education infrastructure and children enrolled in local schools, 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 suggest local governments use the increase in taxes thanks to the program to provide more education inputs 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.

This paper's theoretical framework and empirical evidence speak directly to debates about the right form of decentralization. The existence of a large 'fiscal gap' between

the expenditures and the tax collection of local governments is an ubiquitous characteristic of local governments around the world. In developing countries in particular local governments have recently been granted substantial expenditure responsibilities but local tax capacity generally lags behind. My results suggests that narrowing this fiscal gap by empowering local governments to levy more tax revenue is likely to make them more responsive to local needs. Substantial local tax collection - complemented by intergovernmental transfers for revenue equalization purposes - may be a necessary feature of a successful decentralization that strives to bring government 'closer to the people'.

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 central recommendation of the 2005 report on achievement of the Millenium Development Goals is that developing countries should mobilize increased domestic resources by up to four percentage points by 2015 (UnitedNations (2005)), yet there is very little research on how this aim could be achieved. What's more, technical aid on public sector financial management has always been the poor parent of official development aid[34] 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. It suggests more widely that technical help in tax capacity building may lead to an increase in public resources which is potentially more conducive to the type of public spending that benefits citizens than traditional development aid. To understand whether increasing tax capacity at the national level will upgrade the quality and accountability of national public spending remains an important topic for future research.

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

References

Jose Roberto Afonso and Erika Araujo. Local government organization and finance : Brazil. In Anwar Shah, editor, *Local Governance in Developing Countries*. World Bank, 2006.

Jose Roberto Afonso and Jose Serra. Fiscal federalism brazilian style : Reflections. Paper presented at a Forum of Federations, 1999.

World Bank. Making services work for poor people. world development report 2004. Washington, DC: World Bank, 2004.

Pranab Bardhan and Dilip Mookherjee. *Decentralization and Local Governance in Developing Countries*. MIT Press, 2006.

Timothy Besley. *Principled Agents? The Political Economy of Good Government*. Oxford University Press, 2006.

Timothy Besley and Torsten Persson. The origins of state capacity: Property rights, taxation, and politics. *American Economic Review*, 99(4):1218–44, September 2009.

Timothy Besley and Michael Smart. Fiscal restraints and voter welfare. *Journal of Public Economics*, 91(3-4):755–773, April 2007.

Richard M. Bird and Michael Smart. Intergovernmental fiscal transfers: International lessons for developing countries. *World Development*, 30(6):899 – 912, 2002. ISSN 0305-750X. doi: DOI: 10.1016/S0305-750X(02)00016-5.

BNDES. *Modernizacao da Gestao Publica : uma Avaliacao de Experiencias Inovadoras*. BNDES, 2002.

Deborah Brautigam, Odd-Helge Fjeldstad, and Mick Moore, editors. *Taxation and State-Building in Developing Countries*. Cambridge University Press, 2008.

Fernanda Brollo, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini. The political resource curse. (15705), January 2010.

Francesco Caselli and Guy Michaels. Do oil windfalls improve living standards? evidence from brazil. NBER Working Papers 15550, National Bureau of Economic Research, Inc, December 2009.

Pedro Humberto Bruno de Carvalho Jr. Iptu no brasil: Progressividade, arrecadao e aspectos extra-fiscais. Discussion Papers 1251, Instituto de Pesquisa Econmica Aplicada - IPEA, December 2006.

Claudio Ferraz and Frederico Finan. Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes. *The Quarterly Journal of Economics*, 123(2):703–745, 05 2008.

Claudio Ferraz and Frederico Finan. Electoral accountability and corruption: Evidence from the audits of local governments. NBER Working Papers 14937, National Bureau of Economic Research, Inc, April 2009.

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

Claudio Ferraz and Joana Monteiro. Does oil make leaders unaccountable? evidence from brazil's offshore oil boom. Paper presented at the NEUDC Conference, November 2010., 2010.

Claudio Ferraz, Frederico Finan, and Diana B. Moreira. Corrupting learning:evidence from missing federal education funds in brazil. PUC TEXTO PARA DISCUSSO Nb 562, 2009.

Sebastian Galiani, Paul Gertler, and Ernesto Schargrodsky. Water for life: The impact of the privatization of water services on child mortality. *Journal of Political Economy*, 113(1):83–120, February 2005.

James J. Heckman. *Labor Supply and the Economic Recovery Tax Act of 1981: Comment*, pages pp. 32–38. University of Chicago Press, 1996.

James J Heckman, Hidehiko Ichimura, and Petra Todd. Matching as an econometric evaluation estimator. *Review of Economic Studies*, 65(2):261–94, April 1998.

Keisuke Hirano and Guido Imbens. 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, 2001.

Keisuke Hirano, Guido W. Imbens, and Geert Ridder. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189, 07 2003.

Atsushi Inoue and Gary Solon. Two-sample instrumental variables estimators. *The Review of Economics and Statistics*, 92(3):557–561, 2010.

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

Peter H. Lindert. Why the welfare state looks like a free lunch. NBER Working Papers 9869, National Bureau of Economic Research, Inc, July 2003. URL `http://ideas.repec.org/p/nbr/nberwo/9869.html`.

Stephan Litschig. Intergovernmental transfers and elementary education: Quasi-experimental evidence from brazil. *UPF Working Paper Nb 1142*, 2008a.

Stephan Litschig. Rules vs. political discretion: evidence from constitutionally guaranteed transfers to local governments in brazil. Economics Working Papers 1144, Department of Economics and Business, Universitat Pompeu Fabra, August 2008b.

Stephan Litschig and Yves Zamboni. The short arm of the law: judicial institutions and local governance in brazil. Economics Working Papers 1143, Department of Economics and Business, Universitat Pompeu Fabra, May 2008.

Stephan Litschig and Yves Zamboni. Audit risk and rent extraction: evidence from a randomized evaluation in brazil. Economics Working Papers 1270, Department of Economics and Business, Universitat Pompeu Fabra, April 2011.

Monica Martinez-Bravo, Gerard Padr i Miquel, Nancy Qian, and Yang Yao. Do local elections in non-democracies increase accountability? evidence from rural china. NBER Working Papers 16948, National Bureau of Economic Research, Inc, April 2011.

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

Mick Moore. How does taxation affect the quality of governance? *Tax Notes International*, 47:79, 2007.

OECD. Do no harm : International support for statebuilding. Technical report, Organization for Economic Cooperation and Development, 2010a.

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

Benjamin A. Olken. Monitoring corruption: Evidence from a field experiment in indonesia. NBER Working Papers 11753, National Bureau of Economic Research, Inc, November 2005. URL `http://ideas.repec.org/p/nbr/nberwo/11753.html`.

Dina Pomeranz. No taxation without information. Harvard Economics Department Job Market Paper, 2010.

Ritva Reinikka and Jakob Svensson. Fighting corruption to improve schooling: Evidence from a newspaper campaign in uganda. *Journal of the European Economic Association*, 3(2-3):259–267, 04/05 2005.

Susan Rose-Ackerman. *Corruption and Government : Causes, Consequences and Reform*. New York: Cambridge University Press, 1999.

Gustavo A. G. Dos Santos, F. H. Tavares, L.O. Reiff, and Leticia Correa. Potencial de arrecadaco municipal e o pmat. *Revista do BNDES*, 29:399–434, 2008.

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

Charles M. Tiebout. A pure theory of local expenditures. *Journal of Political Economy*, 64:416, 1956. URL `http://ideas.repec.org/a/ucp/jpolec/v64y1956p416.html`.

UnitedNations. Investing in development : A practical plan to achieve the millenium development goals. Technical report, UN Millenium Development Project, 2005.

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

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

A Theoretical appendix

A.1 Equilibrium with information asymmetries

The representative citizen's maximizes :

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

subject to the constraints

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

where I am using the fact that $G_H = G_L + T_H - T_L + Z(\sigma_H - \sigma_L)$ and λ_i is the lagrange multiplicator associated with constraint i.
The first order conditions for maximization are

$$\frac{\partial W}{\partial \sigma_H} = Z \tag{20}$$

$$\frac{\partial W}{\partial \sigma_L} = Z(\lambda_1 + \lambda_2 - 1 + q) \tag{21}$$

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

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

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

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 $\frac{\partial W}{\partial \sigma_L} = Zq > 0$ and $\sigma_L = 1$. Note also that the participation constraints for both types H and L (constraints 1 and 2) cannot bind simultaneously, so that we cannot have $\lambda_1 > 0$ and $\lambda_2 > 0$ and equation (22) shows that one of them has to bind (otherwise public good in one of the states could be increased whilst keeping taxes and the public good in the other state constant).
It is easy to show that the participation constraint for the high type H cannot be binding. Suppose it is, so that $G_L = (\bar{F}(1+u) + T_L)(1-\alpha) + \alpha(T_L + T_H) + \sigma_L Z$, $\lambda_2 = 1$ and $\lambda_1 = 0$. For the low type's participation constraint to be respected , T_H and T_L must be so that :

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

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

Intuitively, for the high type's participation constraint to bind whilst still respecting the low type's participation constraint it must be that $T_H^* > T_L^*$ so that the low type does not find it profitable to mimic the high type's behavior. However $\lambda_2 = 1$ and $\lambda_1 = 0$ implies that

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

and

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

$$\Leftrightarrow T_H^* = h((1-q-\alpha)/\phi(1-q)),\ \lambda_3 = 0,\ q < 1-\alpha \text{ or } \lambda_3 > 0, T_H^* = 0 \tag{29}$$

which implies that $T_H^* < T_L^*$ and violates (26).

The equilibrium level of public good is thus $G_L^* = (\bar{F}(1-u) + T_L^*)(1-\alpha) + Z$ in the low state and $G_H^* = (\bar{F}(1-u) + T_H^*)(1-\alpha) + \alpha(T_H^* - T_L^*)Z$ in the high state. Clearly the public good is increasing one for one with T_H^* in the high state, and unaffected by T_H^* in the low state. Formally

$$\lambda_3 = (1-q)(\phi C_{T_H} - 1) \Leftrightarrow T_H^* = h(1/\phi),\ \lambda_3 = 0 \tag{30}$$

The equilibrium level of taxes in the low state L is necessarily lower because an increase in T_L^* lowers the public good delivered by the incumbent in the high state, as it makes mimicking the low state equilibrium more attractive. When $q > \alpha$ a marginal increase in T_L^* has a marginal benefit of increasing the public good in the low state by $1-\alpha$ with probability q and a marginal cost of decreasing the public good in the high state by α with probability $(1-q)$. When the probability of the low state occurring is low, $q \leq \alpha$, the marginal cost outweighs the marginal benefit and it is optimal for the citizen to provide no taxes in the low state in order to maximize public good provision in the more frequent high state. Formally

$$\lambda_4 = q(\phi C_{T_L} - 1) + \alpha \Leftrightarrow \text{ when } q > \alpha\ T_L^* = h((q-\alpha)/\phi q),\ T_L^* = 0 \text{ otherwise} \tag{31}$$

A.2 Impact of the tax capacity program

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

$$\frac{\partial T_H^*}{\partial \phi} = -\frac{h'(1/\phi)}{\phi^2} \tag{32}$$

and

$$\frac{\partial T_L^*}{\partial \phi} = \min\{0, -\frac{h'(q-\alpha/\phi q)}{\phi^2}\frac{q-\alpha}{q}\} \tag{33}$$

I write

$$\frac{\partial T^*}{\partial \phi} = -\omega_1 < 0 \tag{34}$$

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

$$\frac{\partial T_H^*}{\partial \phi} - \frac{\partial T_L^*}{\partial \phi} = -\frac{h'(1/\phi)}{\phi^2} - \min\{0, -\frac{h'(q-\alpha/\phi q)}{\phi^2}\frac{q-\alpha}{q}\} = -\omega_2 < 0 \qquad (35)$$

Because $E(G^*) = (E(T^*) + \bar{F}((1-u))(1-\alpha) + Z + (1-q)\alpha(T_H^* - T_L^*)$ it follows from equations (32), (33) and (35) that the program leads to an increase in public good provision :

$$-\frac{\partial E(G^*)}{\partial \phi} = (1-\alpha)\omega_1 + (1-q)\alpha\omega_2 > (1-\alpha)\omega_1 \qquad (36)$$

The fact that program decreases the share s^* of rents in total revenues can be seen by noting that $E(S^*) = E(T^*) + E(F) - E(G^*)$, which implies that $\frac{\partial S^*}{\partial \phi} < \frac{\partial G^*}{\partial \phi}$.

Finally, consider the impact of an increase in the average level of transfers $E(F)$ by an amount $dE(T) = -\omega_1 d\phi > 0$. Note first that this means that $\bar{F}$ increases by $-\omega_1/(1+u(1-2q))d\phi$. The impact on the average value of the public good is

$$-\frac{\partial E(G^*)}{\partial E(F)}dE(T) = (1-\alpha)\omega_1\frac{1-u}{1+u(1-2q)} < (1-\alpha)\omega_1 < -\frac{\partial E(G^*)}{\partial \phi} \qquad (37)$$

which completes the proof.

B 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 $FPM^s_{i,t}$ received by government i in state s is

$$FPM^s_{i,t} = \frac{f(pop_{i,t})}{\sum_{j \in s} f(pop_{j,t})} FPM^s \tag{38}$$

where $f(pop_{i,t})$ is the coefficient corresponding to the population bracket in which the local government's population is found. Table 10 presents the population brackets and associated coefficients[35] in its first two columns. 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). I construct the amounts of theoretical FPM grants each municipality is allocated according to the above rule depending on its state and population size for each year. Table 10 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 and state do not perfectly predict the FPM grants each municipalities receives, due to several reasons. Various law amendments during the 1990s froze the FPM allocations for some municipalities (in particular, the ones that split over the period). Even for municipalities not affected by those amendments the rule is not perfectly enforced : Litschig (2008b) presents some evidence of manipulative sorting above the FPM thresholds for the years 1989 and 1991, evident from the official TCU population estimates used to calculate the allocation of FPM resources each year. The official TCU estimates and the IBGE population estimates used in this paper indeed do not coincide, suggesting potential manipulation at the TCU level. Nonetheless, real FPM grants received do increase substantially at each population threshold. Figure 3 displays the scaterplot of received and theoretical FPM transfers received by municipalities in the state of Minas Gerais in 2008[36]; the vertical lines represent the population thresholds. Both figures display visible jumps at the thresholds, though there are cases of misassignment around the cutoffs in the graph for actual FPM transfers.

To identify the causal impact of an increase in FPM transfers on local spending outcomes I use variations in the amounts of theoretical grants municipalities should

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

[36] The sample is restricted to one year and state (Minas Gerais, which contains the most municipalities) to limit the variation in grants received and make the graphs easier to understand. All year/state combinations provide similar graphs.

Figure 3: Real and Theoretical FPM Transfers for the state of Minas Gerais in 2008

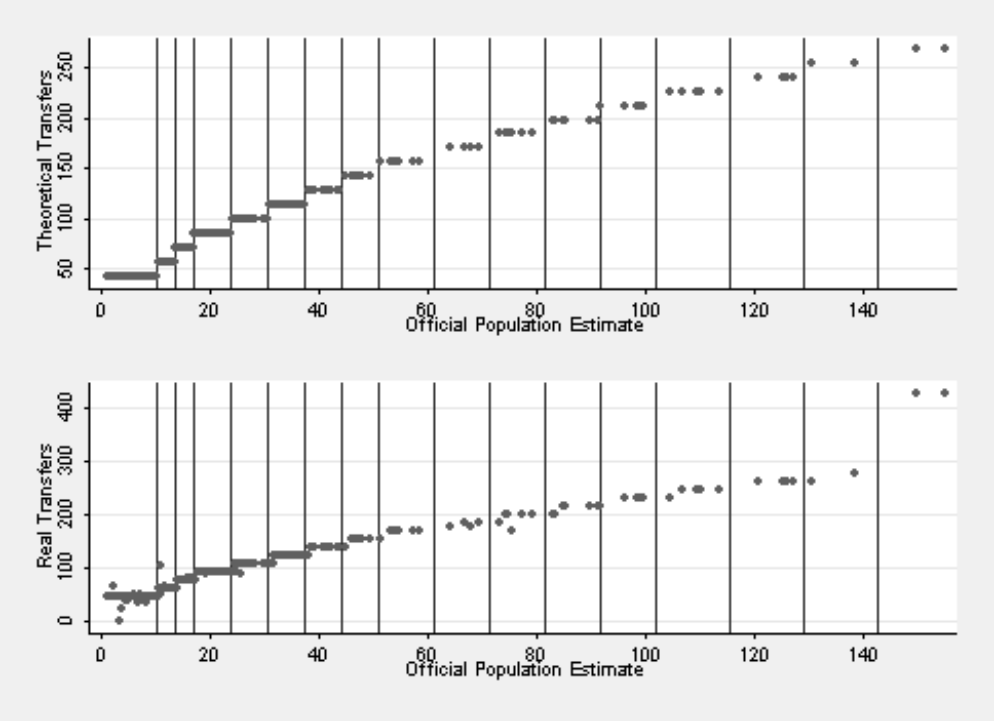


Table 10: Real and Theoretical FPM Transfers and Coefficients

Population	Coefficient	Real Transfer	Theoretical Transfer
<10,189	0.6	22.5	21.8
10,189-13,584	0.8	31.2	30.5
13,585-16,980	1	39	39.1
16,981-23,772	1.2	46.9	47
23,773-30,564	1.4	53.5	54.6
30,565-37,356	1.6	62.2	63.7
37,357-44,148	1.8	69.3	71.7
44,149-50,940	2	78.5	80.6
50,941-61,128	2.2	84.7	87.8
61,129-71,316	2.4	90.9	94.8
71,317-81,504	2.6	99.3	103.7
81,505-91,692	2.8	102.8	107.5
91,693-101,880	3	117.9	125.6
101,881-115,464	3.2	127.8	134.8
115,465-129,047	3.4	132.6	135.3
129,048-142,632	3.6	144.4	146.1
142,633- 156,216	3.8	209.1	166.1
≥156,217	4	350.7	349.2

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 from the *FINBRA* database.

have received, controlling for any impact of the variables determining the allocation by using state fixed effects and a high-order polynomial in population size. Table 11 presents the first stage of this identification strategy. We see that the actual amount of FPM transfers received increases one for one with the theoretical amounts. None of the control variables have any impact on the amounts received, suggesting that manipulations of the rule are rare and/or unrelated to the variables which affect the dependent variables of interest in this paper.

Table 11: First Stage Regression for FPM Transfers

	Dep. var : Real FPM Transfer pc
Theoretical FPM pc	1.046***
	(0.074)
GDP pc	0.000
	(0.001)
Population	0.550
	(2.815)
Agr\ GDP	-40.937
	(36.303)
Serv\ GDP	2.319
	(32.428)
Density	0.003*
	(0.002)
Income pc	0.003
	(0.003)
Inequality	12.010
	(12.195)
[1em] Education	-9.097
	(6.218)
Pol. competition	20.313
	(15.806)
Term limit	-4.982
	(4.837)
Observations	48460
Municipalities	5087

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The regression includes state and year fixed effects as well as a high order polynomial in population size.

C Estimation of the Propensity Score

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 2. 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

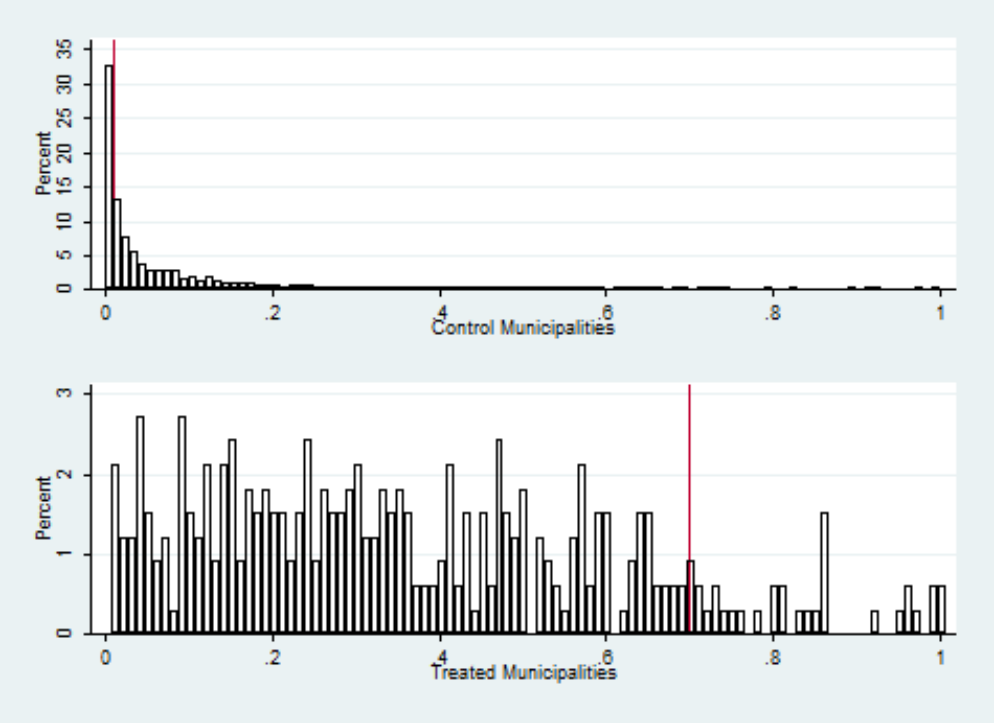
	eq1
=1 if PMAT pgm wasn't cancelled, 0 if it was	
Income	0.1586**
	(0.0757)
Population	0.1069**
	(0.0468)
Taxes in 1998	0.0000
	(0.0005)
Agr\ GDP	-0.0048
	(0.0042)
Serv\ GDP	0.0006
	(0.0045)
Education	0.0203
	(0.0875)
Urban pop.	0.8583***
	(0.2942)
Inequality	-0.8831
	(0.9148)
Governor's party	-0.0827
	(0.1073)
Pol. competition	0.7736**
	(0.3786)
esample() from estimates store	
Distance to closest PMAT	
Observations	3560
Marginal effects; Standard errors in parentheses	
(d) for discrete change of dummy variable from 0 to 1	
* p¡0.10, ** p¡0.05, *** p¡0.01	

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 4 graphs the distribution of the propensity score in the treated and control groups. The red lines indicate the limit of the common support.

Figure 4: Distribution of the propensity score



D Table Appendix

Table 13: Impact of the Program, Alternative Specifications : (1) All sample (2) Logs

	Taxes	Taxes	Infrastructure	Infrastructure	Quality	Quality	Corruption	Corruption
1th year	5.138***	0.048**	0.066**	0.022*	0.178***	0.122***		
	(1.716)	(0.024)	(0.032)	(0.012)	(0.036)	(0.037)		
2th year	6.738***	0.097**	0.102***	0.038**	0.192***	0.079*		
	(2.061)	(0.039)	(0.036)	(0.018)	(0.068)	(0.044)		
3th year	10.042***	0.099***	0.138***	0.042**	0.440***	0.069		
	(2.294)	(0.035)	(0.042)	(0.017)	(0.118)	(0.051)		
4th year	12.668***	0.087***	0.193***	0.061***	0.293***	0.097		
	(2.685)	(0.031)	(0.047)	(0.020)	(0.068)	(0.059)		
5th year	16.251***	0.063**	0.264***	0.100***	0.432***	0.095		
	(3.357)	(0.031)	(0.058)	(0.034)	(0.092)	(0.062)		
All years							-56.282	-0.467**
							(42.152)	(0.216)
Observations	35562	24049	35569	24047	35569	8992	705	483
Clusters	3654	2462	3654	2462	3654	1444	705	483

Standard errors in parentheses
* p¡0.10, ** p¡0.05, *** p¡0.01

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 equations (13) and (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 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.

Figure 5: Distribution of Application and Start Dates

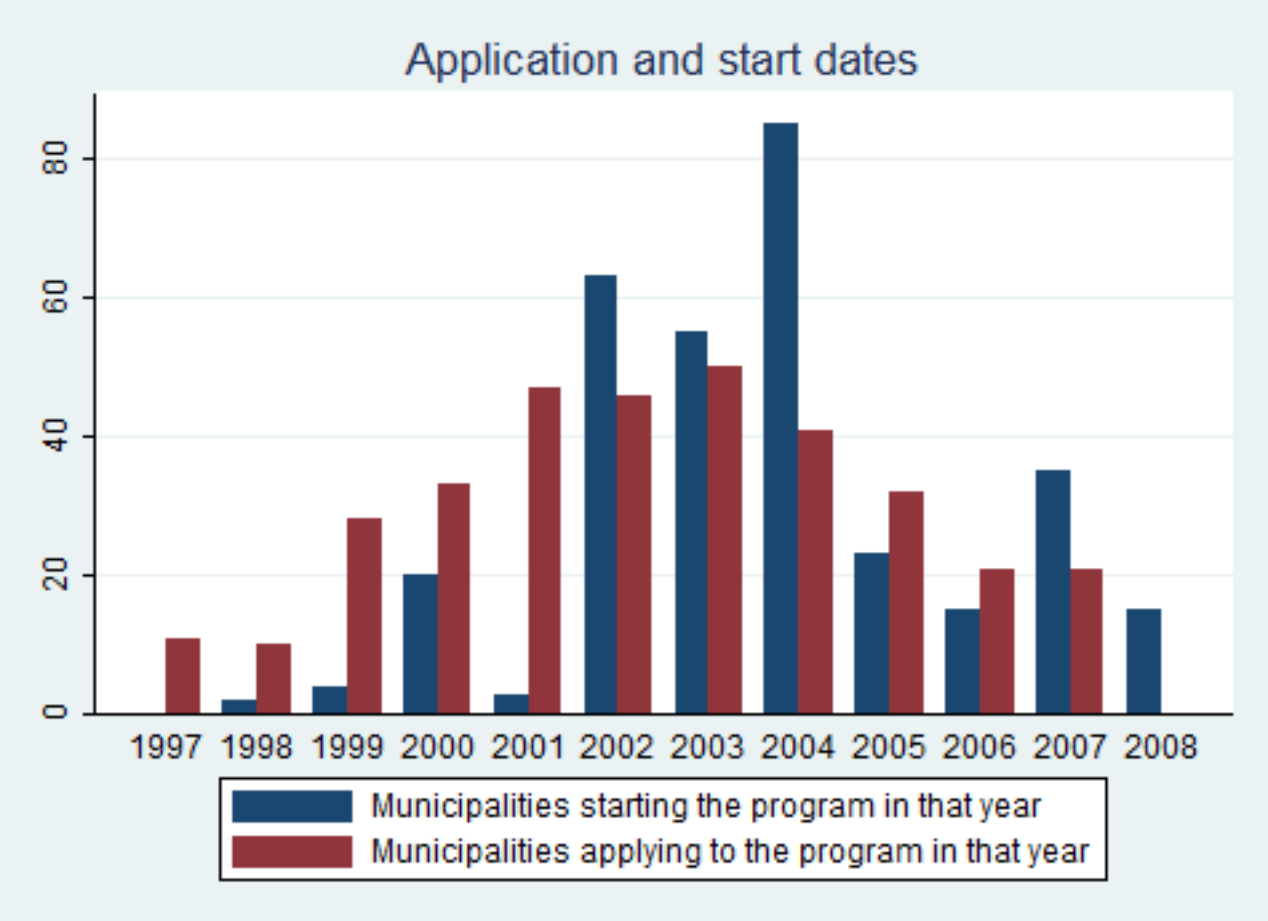


Table 15: Within municipality mean (standard deviation) in taxes and transfers, unweighted sample

	All	Controls	Treated before PMAT	Treated after PMAT
Taxes	70.1 (74)	65.1 (78)	90 (19)	142.9 (25)
[1em] Transfers	174.8 (48)	182 (77)	106 (17)	129 (22)