Accountability and yardstick competition in the public provision of education

Rafael Terra [*]
University of Brasília

Enlinson Mattos [†]
São Paulo School of Economics

November 15, 2015

Abstract

This paper explores the institutional change introduced by the public disclosure of an education development index (IDEB, Basic Education Development Index) in 2007 to identify the effect of education accountability on yardstick competition in educational spending for Brazilian municipalities. A preliminary analysis of the data show that spatial strategic behavior on educational spending is estimated to be lower for lame ducks and for those incumbents with majority support on the city council. This suggests a strong relation between commitment and accountability, which reinforces the yardstick competition theory. Second, we find a minor reduction (20%) in spatial interaction for public educational spending after the IDEB's disclosure — compared with the spatial correlation before the disclosure of the index. Our main results explore the discontinuity of the IDEB's disclosure rule around the cutoff of 30 students enrolled in the grade under assessment. The estimates suggest the spatial autocorrelation — and thus the yardstick competition — is reduced in 52%. Falsification and robustness tests were performed and suggest we can claim causality around small bandwidths from the cutoff. This finding suggests the public release of information may decrease the importance of the neighbors' counterpart information about voters' decisions.

Keywords: education spending; yardstick competition; electoral and educational accountability.

JEL classification: C21; H72; H73.

[*]Address: University of Brasília, Economics Department, FACE,Campus Darcy Ribeiro, Brasília, Brazil, 70910-900 , telephone: 55-61-3107-0807, e-mail: *rterra@unb.br*.

[†]Corresponding Author. Address: São Paulo School of Economics, Getúlio Vargas Foundation 474 Itapeva St., 01332-000, São Paulo, Brazil, telephone: 55-11-3799-3350, e-mail: *Enlinson.Mattos@fgv.br*.

1 Introduction

Information asymmetry between voters and politicians is known to be a building block for the well-established models of political agency.[1] In this framework, voters do their best to find ways to improve their information about the incumbent. One possibility for determining the quality of incumbents (agents) consists of voters (principals) evaluating the incumbents' performance in terms of tax levels and amount (and/or quality) of public services by comparing them to those of neighboring jurisdictions, where information can be accessed more easily. In turn, the better-informed part—i.e., the incumbents—then engage in a yardstick competition to signal their performance to the voters (Salmon 1987).

A large body of empirical literature interested in testing the nature of strategic interactions between jurisdictions (both in cases of expenditures and tax-setting) has been produced.[2] Most of these studies relate the degree of yardstick competition to a range of political incentives (Bordignon et al. 2004, Geys 2006, Allers and Elhorst 2005). However, exogenous changes to information asymmetry are much less explored and to the best of our knowledge, no empirical study has explicitly investigated yardstick competition in the context of education expenditures [3] or verified the effect of reinforcing educational accountability on the strategic choice of this type of expenditure. Revelli (2006) is the first one to exploit an institutional reform taking place in the UK to address the fact that the spatial pattern observed in welfare policy is at least partially driven by yardstick competition. [4] Our work contributes to the literature by providing additional evidence on the role of the national disclosure of local governments' performance ratings in the relative quality of public services' provision.

This paper uses data from Brazilian municipalities to test whether the mid-2007 local-level disclosure of the Brazilian Basic Education Development Index (IDEB, in its Portuguese

[1]See Ferejohn (1986), Alesina and Cukierman (1990), and Persson et al. (1997).

[2]For studies on strategic spending-setting, see Revelli (2006), Elhorst and Fréret (2009), Case et al. (1993) and Bivand and Szymanski (2000). For studies on strategic tax-setting, see Besley and Case (1995), Bordignon et al. (2003), Revelli (2008), Allers and Elhorst (2005), Sollé Ollé (2003), Revelli (2002), Ladd (1992) and Revelli (2001).Other dimensions of public policy can be of interest to voters. Geys (2006) notes that voters may care about efficiency in the production of local public services, i.e., about their level of public services in light of the taxes that they pay. Rincke (2009), conversely, observes that voters may also value the adoption of new technologies for the provision of public services, thus also evaluating the incumbents' relative performance in terms of their innovative ability.

[3]Although Rincke (2009) relates the adoption of educational innovations to yardstick competition, that author overlooks the phenomenon in the expenditure setting.

[4]Similar effects can be found by changing the rules for the concession to operate a public service (Bivand and Szymanski 2000).

acronym) diminished the spatial interaction among jurisdictions in terms of educational spending. Such a reduction in spatial interaction could be attributed to reduced information asymmetries regarding the quality of education after IDEB became public information. This papers' main strategy consists of estimating the effect of IDEB disclosure on yardstick competition exploring a rule defined by the Ministry of Education that prevents schools with fewer than 30 students enrolled in the grade under evaluation from participating in *Prova Brasil*, the exam that measures the proficiency levels that comprise the IDEB. [5] This approach admits that IDEB disclosure may not be exogenous because jurisdictions are free to decide whether to participate (although the vast majority do participate). Thus, we estimate local regressions that resemble a regression discontinuity design, in which our interest lies mainly in shifts in the spatial correlation coefficient around the cutoff of 30 students instead of in the intercept. [6]

The usual framework in the literature is the political agency-model, in which voters (principals) are not aware of the true costs of providing public services and are imperfectly informed about the quality of incumbents (agents). Besley and Case (1995) argue that because of this information asymmetry, the voters can mistake the incumbents' attempt at rent appropriation for negative economic shocks, which makes it difficult to distinguish the "good" from the "bad-type" incumbents, i.e., those that will or will not attempt to charge rent on top of the cost of providing public services. Thus, the incumbents that are considered the bad type, if they are willing to seek reelection, should not set taxes to a point at which it becomes evident to the voters that they are attempting to charge rent. To evaluate the incumbent's performance (or the incumbent's type), the voters then compare their own levels (or quality) of public services and/or taxes with those of the neighboring jurisdictions, where information is more easily accessible from the media or by other means.[7][8] To signal their performance to voters, the perfectly informed incumbents then engage in competition with the neighboring jurisdictions by mimicking each other's fiscal behavior.[9] Finally, if yardstick

[5]In fact, even schools that participated in the exam, but where fewer than 30 students attended on the day of the exam, did not have their IDEBs disclosed.

[6]Differently from traditional RDD we parametrize our model's functional form with a spatial lag that has to be instrumented with average values of municipalities neighbors' regressors.

[7]See Ferrazt and Finan (2008), Strömberg (2004) and Revelli (2008) on the role of the media in providing voters with information.

[8]Salmon (1987) and Case et al. (1993) stress that voters and incumbents do not necessarily need to compare their jurisdiction's performance to that of neighboring jurisdictions. Instead, the comparison can occur between similar jurisdictions, where similarity is defined in terms of a wide range of characteristics such as population, income and ethnic composition, to name a few.

[9]The same conclusions can be reached using other frameworks. Revelli and Tovmo (2007), for example, rely on a bureaucracy agency-model with welfare-maximizing politicians (principals) and self-interested bureaucrats (agents). In such a model, information about the true cost of providing public services is asymmetric and to achieve efficiency, the principals compare their own production of public service with that

competition improves voters' power to discipline politicians and to make bad incumbents willing to pool with good ones, it can be shown that such competition will be welfare enhancing compared to a situation in which voters ignore the fiscal performance of neighboring jurisdictions.

Conversely, Bordignon et al. (2004) warn that yardstick competition might not necessarily lead to greater interaction among jurisdictions. The existence of yardstick competition can dampen the incentives for bad incumbents to pool with good ones, i.e., to choose a level of rent that is not so high as to allow voters to perfectly identify bad incumbents. Bad incumbents could prefer to extract the maximum amount of rent during their first term and then be voted out of office rather than mimic good incumbents' behavior to increase the odds of an uncertain reelection (and only then divert the maximum amount of rent). Thus, yardstick competition can actually decrease the amount of strategic interaction among local governments.

Besley and Smart (2007) also observe that yardstick competition can be welfare diminishing when compared to a situation in which the voters ignore their neighbors' fiscal performance. When the voters know both the reputation of the neighbors' incumbents and their fiscal situation, it becomes harder for the bad incumbents to hide their type, thus inducing them to extract the maximum amount of rent while in office.

Other political motivation may be crucial in determining strategic interaction. For example, in the event that officials are not running for reelection by force of law, i.e., they are lame ducks, voting will no longer enforce discipline, and some will set the level of taxes and expenditures that maximizes rent extraction. Thus, the lame-duck incumbents in principle should not have incentives to use their neighbors' performance as a benchmark (Besley and Case 1995). The same reasoning can be applied to both incumbents on the verge of retirement and incumbents whose parties have decided not to run for reelection. As argued by Alesina and Spear (1988), it might be the case that the lame ducks have some partisan interest that prevents them from attempting the maximum rent extraction, but the expected amount of spatial interaction should still be smaller. Considering that holding the majority of seats on the city council implies having the support of the majority of voters, the size of the majority can also change the pattern of interaction among jurisdictions (Allers and Elhorst 2005, Elhorst and Fréret 2009). Other factors that can induce changes in the interactions among jurisdictions are the votes received in the most recent election (Besley and Case 1995, Sollé Ollé 2003), the existence of coalitions (Geys 2006) and ideology (Allers and

observed in neighboring jurisdictions.

Elhorst 2005, Sollé Ollé 2003). These political and institutional features are necessary to identify yardstick competition because the presence of fiscal spatial interaction itself may reflect competing phenomena such as tax or welfare competition (Brueckner 2003). [10]

A preliminary analysis of our data shows that political motivation seems to be pervasive in setting educational expenditures. Incumbents with majority support in the legislature and who are serving their last term in office are less engaged in strategic interaction. That is evidence that yardstick competition is present in educational spending. We also compare spatial correlation on education spending before and after the indexes were disclosed, and find a reduction in spatial interaction. As other confounding factors may have happened concomitantly with the IDEB disclosure, such a "before and after" difference may be biased.

Our main results using the cutoff assignment rule in the number of students to determine the IDEB disclosure indicate higher spatial correlation in education spending between municipalities to the left of the cutoff, i.e., where IDEB was not disclosed. We evaluate the local regressions for bandwidths $h = 5, 6, 7, 10, 20$ and 30. Second order polynoms models are evaluated and results are unchanged. McCrary tests show no evidence of cutoff's manipulation. Regression Discontinuity Designs are estimate on covariates, and we find no evidence of jumps in the covariates. Finally, falsification tests are performed and corroborate causal relationship for the smaller bandwidths $h = 5, 6, 7$ and 10.

In principle, the effects of disclosing standardized test results on spending interaction patterns are unclear. The relationship between educational spending and education quality (measured by student achievement) remains largely unknown to voters, public officials and even academics. As noted by Hanushek (1986, 1996, 2006), the lack of information about the educational production function causes officials to employ financial resources for inputs that have little or no role in determining educational output. Moreover, the officials' objective is not necessarily to be efficient in educational matters. Unsurprisingly, educational spending and students' performance do not necessarily go together.[11] Thus, the effect of students'

[10]See Brueckner (2003). A tax reduction on capital, for example, raises the net-of-tax return on capital. As a mobile factor within a limited geographic area, there will be an inflow of capital to the jurisdiction to equalize the net-of-tax return. To avoid losing the tax base, the other jurisdictions engage in a local "race to the bottom" with respect to taxes, which reflects a tax-mimicking process and greater spatial autocorrelation. Finally, if there is a balanced budget, the expenditures will follow the pattern verified by the taxes. Similarly, when a mobile labor force exists within a limited geographic area, an increase in the value of welfare benefits distributed to the poor (occupied in low-skilled jobs) in one jurisdiction will attract unskilled labor from the surrounding areas to equalize the gross income across jurisdictions, which can ruin the program. To keep their welfare programs functioning, local governments must not set their benefits higher than those of their neighbors. This phenomenon also produces benefits-mimicking and spatial autocorrelation in expenditures.

[11]See Menezes-Filho and Pazzello (2007), Card and Payne (2002), Leuven et al. (2007) and Revelli (2009).

performance disclosure on yardstick competition is far from obvious, thus rendering it an empirical matter.

It is likely that prior to the disclosure of student achievement, the incumbents did not pay much attention to educational quality because it was not objectively measured. Officeholders may have changed their attitudes after average student performances were made public at the local level. Indeed, Firpo et al. (2011) find evidence in the context of Brazilian municipalities that higher average achievement increases the odds of incumbents' reelection. Thus, once schools and municipalities' performances were made public, one could ask whether incumbents changed educational spending patterns as though there were a deterministic relationship between students' achievement and expenditures — or if they did nothing because they lacked knowledge about how to effectively improve education quality.

This paper is organized as follows. Section 2 discusses the institutional background of Brazil. Section 3 presents the data set description. Section 4 presents the estimation strategy, and Section 5 shows the results of the spatial models and the robustness tests. Finally, Section 6 concludes.

2 Institutional background

Below, we describe the institutional aspects of Brazil that assist in understanding this paper. First, we briefly describe Brazil's recent accountability experience. Second, we describe the public financing of local governments and third, we describe the political system.

2.1 Educational accountability

Educational accountability is a relatively new concept in Brazil. Only after 1995, with the implementation of the National System of Basic Education Assessment (known by its Portuguese acronym, SAEB), we could track educational quality, but only at the state level. It was only after the creation of IDEB — based on the students' performances measured by *Prova Brasil*— that we could track educational quality at the school or municipality level.

The IDEB is an index that measures the overall quality of education in schools and municipalities in an intelligible and direct manner. The index is defined as $IDEB_{ijt} = P_{ijt}A_{ijt}$, where P_{ijt} stands for the average performance in the math and reading exams of

Prova Brasil in unit i in stage of education j in period t.[12] The term A_{ijt} reflects the school's pass rate and varies between 0 and 100%.[13] The index has been standardized to lie in the interval between 0 and 10, wherein 6 corresponds to the average achievement of OECD students (based on the results of the 2003 edition of PISA).

The index was first released on April 26, 2007 by Presidential Decree n. 6094, which is known as the Plan of Goals "All Committed to Education". This decree established goals for each school and municipality's IDEB. The plan envisaged subnational governments voluntarily signing an agreement in which they would commit themselves to achieve gradually increasing goals. In exchange, the federal government provides the municipalities with technical support and orientation related to the best practices for increasing student achievement. The idea is to encourage society to monitor its accomplishment of the goals, reinforcing a sense of accountability for local educational quality and diminishing the information asymmetry related to the quality of incumbents. The federal government final goal is to achieve an average IDEB of 6.0 by 2021, i.e., the average performance of OECD students. [14]

These goals were the Government's response to the discussion with the civil society represented by the Non Governmental Organization of the same name "All Committed to Education", founded in 2006 with the support of major Brazilian companies — whose budgets together add up to tens of billions of dollars —[15]. The NGO also set similar goals in terms of literacy rate, student attendance, reading and math learning, graduation rate and education investment. The NGO works with state and local governments, helping with the dissemination of good educational practices and monitoring the accomplishment of the goals.

[12] Every two years, the exam assesses the math and the reading skills of 5^{th} and 9^{th} graders (in primary education) at public schools.

[13] Note that there is a tradeoff between the performance and the pass rate. Artificially increasing the pass rates to obtain a higher IDEB will cause less-prepared students to be promoted to the next grade, thus reducing the component of the IDEB that measures performance on standardized exams. The methodology used to build the index (combining achievement and passing rate) intended exactly that—i.e., to simultaneously improve student achievement and lower grade retention.

[14] This spontaneous participation in the "Plan of Goals" fits perfectly with the objectives of this paper. If participation were legally enforced, jurisdictions could still reduce their interactions related to the provision of public education, less because of the incumbents' need to signal their type to the voters and more because non-compliance with the law could result in legal consequences.

[15] "Todos pela EducaÃ§Ã£o" in portuguese. In fact, the Plan of Goals was named after the initiative of the NGO.

2.2 Local Public Finance and Education Funding

Brazil is a federal state that is characterized by the union of 27 states (including the Federal District) and 5,565 municipalities. There is substantial decentralization in the provision of public services. The municipalities are primarily in charge of providing urban sanitation, road conservation, traffic control, health services, land-use regulation, early childhood and primary education (the latter being equivalent to the first 9 years of K-12 education). The states' provision of public services focuses on high school (although in some municipalities, the states also maintain primary schools), higher education, public safety, water provision and sewage collection and treatment. The national government focuses on providing services of broad interest such as social security, energy, defense, higher education and economic-development policies.

Conversely, the power to tax is only weakly decentralized. As of 2011, the majority of municipalities had raised very little revenue through own instruments, amounting to only 6.6% of total revenue. The municipalities' main instruments of taxation are the property tax (1.15% of total revenue), the tax on services (2.97% of total revenue), the payroll tax on employees (1.04% of total revenue), fees for services including, inter alia, garbage collection and street lighting (0.62% of total revenue), the tax on the transmission of property ownership (0.76% of total revenue), and other sources of revenue (0.07% of total revenue).[16]

The municipalities' main sources of revenues are intergovernmental transfers, such as the block grant known as the Municipalities' Participation Fund or FPM (40% of total revenue); the categorical grant for the financing of health services, also known as the Unified Health System or SUS (7.26% of total revenue); the categorical grant for education known as the Fund for the Maintenance and Development of Basic Education and Valuation of Teaching, or FUNDEB (18.07% of total revenue); and 1/4 of all of the state's indirect tax on the circulation of goods and services (also known as ICMS) collected within the municipality's borders (18.14% of total revenue).

Local educational spending is financed by FUNDEB (a categorical grant) and by sources over which the municipalities have discretion; therefore, it can vary according to local demand for education. The discretionary sources come from FPM — which is funded by 22.5% of the total federal income tax and the same percentage of the total federal indirect tax on industrialized products known as IPI —, municipalities' share of the ICMS, and revenues collected through local instruments.

[16]Data are obtained from the National Treasury Office.

FUNDEB's funding scheme is quite complicated in that a different fund is created by each state and several sources make up the funds. It gathers 15% of its revenues from both the FPM and the States' Participation Fund (also known as FPE), 15% of the IPI owed to states, and the same share of the ICMS owed to both the states and the municipalities, among other, less-important sources. This latter source provides the most important contribution to the fund: approximately 60%.

After a state has received all of the resources that comprise FUNDEB, it divides the amount by the number of students to proportionally distribute the money to the municipalities. If the amount per student is inferior to a minimum value that is defined each year by executive act, the federal government complements the state fund to reach this minimum. Before 2007, FUNDEB, then known as the "Fund for the Maintenance and Development of Fundamental Education and Valuation of Teaching", or FUNDEF, targeted primary-school students. From 2007 on, FUNDEB was reformulated to encompass preschool, kindergarten and high school students.[17] The legal minimum amount of funds destined for students changed significantly over the period analyzed. As of 2002, the minimum value to be transferred to students in the first stage (cycle) of primary education was 418 reals (or 118.30 USD), whereas in 2011, this figure was equal $1.722, 05$ reals (or 920.85 USD). The minimum values differ (though not by much) according to the educational stage that the students are enrolled in, whether they attend urban or rural schools and the state they reside.

Despite the importance of the categorical grants for educational financing, the large majority of the municipalities spend considerably more than the amount that they receive in the form of transfers.[18] In 2011, the total educational categorical grants amounted (on average) to 57,76% of the total educational expenditure of the municipalities. This spending in excess of the categorical grants indicates that the demand for education is higher than the grant would allow. To finance this difference, the municipalities rely mainly on revenues from the FPM (the main source) and from the share of ICMS that belongs to the municipality.

Finally, this overview shows the importance of the national and state indirect taxes for financing not only education but also all goods and services that are provided by the local governments. Contrary to many countries, where property tax is crucial for the determination of the local tax price and the demand for public services, in Brazilian municipalities

[17]The reformulation process involved an increase in the amount of resources devoted to the fund. The per-pupil amount, however, may have increased, decreased or stayed the same depending on the number of students in each stage of education, whether they study in rural or urban areas, or whether they study in full-time schools.

[18]In 2011, only 1.04% of the municipalities spent less than the amount received as educational categorical grant.

this tax is of minor importance. Tax price must be mainly a function of the state and the national indirect taxes.

2.3 The Political System and Budget Approval

Brazilian municipalities are governed by an incumbent mayor who is elected for a 4-year term. Since the 2000 election, the incumbents have been allowed to run for a second and final term. The jurisdictions elections are decided by majority rule, with only 1 round where there are less than 200,000 voters and 2 rounds otherwise.

Aldermen are elected for 4-year terms through an open-list proportional system and are not subject to term limits. The same system applies to the state and federal legislatures, ultimately favoring the proliferation of parties. Brazil currently has 29 active parties, and although some of them are identified with some ideology at the time of their founding, once in office, they often must form coalitions with parties of different ideologies to build majorities in the legislature. This process produces inconsistency between the incumbent party's public policies and its ideology. In addition, Desposato (2006) shows that the party-switching rate in the Brazilian chamber of deputies is higher than 40% (on average). Much of the switching can be attributed to the deputies' desire to broaden their access to public funds to finance pork-barrel projects and to increase their odds of reelection. Thus, even though ideology is the driving force for a few parties, it is generally of secondary importance. At the local level, the inconsistency between party ideology and public policies is even more explicit because local governments have a limited capacity to raise revenues; most of their resources come from intergovernmental transfers. This characteristic of local public finance in Brazil favors studies on yardstick competition on spending rather than on tax-setting.

The aldermen are responsible for creating and changing their municipality's Organic Law, legislating on local subjects, and evaluating the budget submitted by the executive. The budget process is enforced by the Fiscal Responsibility Act, which requires local executives to elaborate a 4-year plan of action (a multi-annual plan) with objectives, the units in charge of executing projects, the amount to be spent, the total execution period and revenue sources. The budget process also requires the elaboration and approval of a Budgetary Guidelines Law with goals and priorities for the subsequent fiscal year (beginning in January). The final step consists of the local executive submitting the Annual Budget Law with the detailed revenues and expenditures expected for the next fiscal year to be voted on and approved by the city council by the end of the fiscal year (in December). Thus, the new expenditures usually take a certain amount of time before being executed, which means that after the IDEB was first disclosed in mid-2007, its effects on spending behavior must have been observed only in the following year.

3 Data and Variables

Although our main results are based on 2008 data, we also use Brazil's municipalities data ranging from 2003 to 2011 in the preliminary analysis as well as in the robustness tests. This period was one with little institutional change in the education sector besides the introduction of *Prova Brasil*, IDEB and the "Plan of Goals All Committed to Education". Data are available for 3,723 — of the 5,565 total — municipalities. The variables in the model that identify the spatial lag parameter are described in Table 1, and the descriptive statistics are presented in Table 2. The continuous variables (and indexes) enter the econometric model (see equation 1 below) in their logarithmic form, whereas the proportions and the dummy variables enter the model unchanged.

The educational spending per pupil is made available by the National Treasury Office (STN). Figure 1 shows the remarkable evolution of educational spending over that period. Several factors contributed to this increase. The economic growth and the increasing efficiency of tax collection ultimately increased the available revenue. Additionally, over the last decade, there has been growing concern about investment in basic education.

Table 1: Description of the variables

variable	**Description**	**source**
Dependent Variable (y)		
education spending	Education spending per pupil enrolled at the local public school system per year.	FINBRA-STN
Controls (X)		
IDEB Disclosure Period (T_{0811})	IDEB disclosure variable equal to 1 from 2008 on and 0 otherwise	-
IDEB disclosure(D)	Refers to the year of 2008. It consists of a dummy variable equal to 1 for municipalities that had their IDEB disclosed — in the previous year — and 0 otherwise.	INEP-MEC
gdp	Gross domestic product per capita (net of public sector activity). It is a proxy for total income (unavailable for the period) and for own revenue raising capacity.	IBGE
wage	Average wage of formal sector workers. It is a proxy for total income (unavailable for the period).	RAIS-MTE
occupation	Is given by the following expression $occupation = (occupied_j/total\ pop_j) \times 100$, where $occupied_j$ is the number of individuals between 25 and 65 years old occupied in the formal sector of municipality j, and $total\ pop_j$ is the total individuals of the same age living in the municipality j. It intends to control for the bias resulting from considering only the wage in the formal sector when that is used as a proxy for total income.	RAIS-MTE

Continued on next page

Table 1 – *Continued from previous page*

variable	Description	source
categorical grant	Total grant per pupil received by the municipality with the specific purpose of education financing. Includes FUNDEB (previously FUNDEF) grants as well as any categorical grant targeted at education, such as the ones from intergovernmental agreements and voluntary (non-mandatory) transfers.	FINBRA-STN
block grant	Total grants per capita received through FPM. These general purposes block grants consist of the main source of municipal revenue.	FINBRA-STN
tax price	$tax\ price = 100 \times (collected_j / municipal\ revenue_j)$, where $collected_j$ consists of taxes collected in the municipality$j$ by all levels of government through mostly indirect taxes on final goods and services, and $municipal\ revenue_j$ is the total revenue of the jurisdiction. It consists of a proxy for the real "tax price".	FINBRA-STN and IBGE
schooling	Average years of schooling.	RAIS-MTE
men	Percentage of male individuals.	DATASUS-MS
population	Total population.	IBGE
elderly	Percentage of individuals over 65 years old.	DATASUS-MS
young	Percentage of individuals under 18 years old.	DATASUS-MS
rural	Percentage of the local public schools' students attending schools in the rural area.	Education Census-MEC
second cycle	Percentage of local public schools' students attending the second cycle of fundamental education.	Education Census-MEC
competition	Number of candidates running for office.	TSE
incumbent's age	Age of the incumbent.	TSE
incumbent's education	Dummy variable assuming value of 1 if the incumbent finished higher education and 0 otherwise.	TSE
left	Dummy variable assuming value of 1 if the incumbent belongs to a left wing party and 0 otherwise. The following parties were considered to be left wing (in acronyms): PC do B, PT, PDT, PSTU, PCB, PSB, PCO,PPS, PSOL.	TSE
incumbent women	Dummy variable assuming value of 1 if the incumbent is a woman and 0 otherwise.	TSE
majority of seats	Dummy variable equal to 1 if the incumbent coalition holds more than 50% of the city council's seats.	TSE
percentage of seats	Percentage of seats held by the incumbent coalition at the city council.	TSE
president's party	Dummy variable equal to 1 if the incumbent's party is the same as the president's and 0 otherwise.	TSE
governor's party	Dummy variable equal to 1 if the incumbent's party is the same as the governor's and 0 otherwise.	TSE
lame-duck	Dummy variable equal to 1 if the incumbent is in his or her second and final term and 0 otherwise.	TSE
aldermen's education	Percentage of aldermen with higher education.	TSE
aldermen's age	Average age of the aldermen.	TSE
women in council	Percentage of women in city council.	TSE
competition for seats	Ratio of the number of candidates to the number of seats available at the city council.	TSE
fragmentation	It is calculated by the following formula:$fragmentation = 100 \times (1 - \sum_{i=1}^{N} p_i^2)$, where p_i is the share of seats held by each party i at the city council.	TSE
Other		
x	Number of students in the municipality.	Education Census-MEC

Continued on next page

Table 1 – *Continued from previous page*

variable	Description	source
More than 30 students (z)	It is the excluded instrument that identifies the endogenous variable of IDEB disclosure D. It equals 1 whenever the number of students exceeds 30, and 0 otherwise.	Education Census-MEC

Notes: All monetary variables are measured in reais ($R\$$).
In the econometric models, the continuous variables (and indexes) enter in their logarithmic form, whereas the proportions and the dummy variables enter the model unchanged.

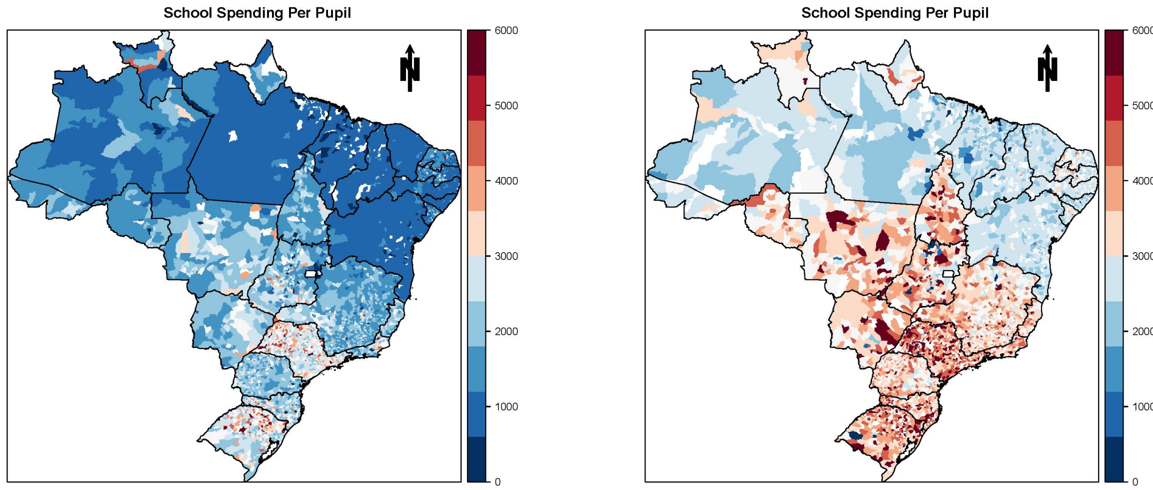
On the left panel of figure 1, for 2003, the spatial patterns are less clear due to the scale convention. In the right panel, however, for 2011, there are clearer spatial patterns and a striking difference in the levels of per-pupil spending between northern and southern municipalities.

Table 2 shows the descriptive statistics. There, we can see that the mean spending per pupil over the period between 2003 and 2011 was 3,157 reals (1,814 USD).[19] Note also that the standard deviation is expressive (2,094 reals or 1,203 USD), evidencing the large difference in educational spending among Brazil's local governments.

Categorical and block grants are expected to have a positive impact on the level of educational spending. For numerous municipalities with a low revenue-generating capacity, education grants are supposed to have a more pronounced impact because they cannot serve as substitutes for own revenue in financing other activities. Block grants, in turn, need not necessarily be employed for education; the amount directed to this area depends on the marginal propensity to spend on education, regardless of the fiscal capacity. Table 2 shows that the mean value of the categorical grants is 1,558 reals per pupil (or 895 USD). The block grant received by the municipalities amounts to an average of 597 reals per capita (or 343 USD). The variation in the grants' amounts among localities is also very expressive, as the standard deviations make clear.

The "tax price" is an important variable in the public finance literature and usually reflects the share of local property taxes paid by the representative voter. However, property tax is a minor source of local revenue in Brazil. The most significant portion of revenues comes from block and categorical grants, — which are primarily funded through taxes such as the ICMS and the IPI — and from the direct participation of the municipality in the total ICMS revenue. The "tax price" that we calculate for this paper takes these specificities into account by considering the ratio between the sum of the main taxes collected within the municipality's borders — whose main components are the state and federal indirect taxes known, respectively, as ICMS and IPI — and the total revenue of the jurisdiction. This

[19]Prices are adjusted by the Amplified Consumer Price Index (IPCA) to the prices of December 2009. The exchange rate also refers to December 2009



(a) 2003 (b) 2011

Figure 1: Local level education spending per pupil

variable attempts to measure the cost of providing one monetary unit of public services that accrues to the local citizens. As observed in Table 2, the mean tax price is equal to 59.65%, which means that most of the municipalities are net recipients of public funds. The higher this ratio, the less expenditure on education (or any other public service) the citizens are expected to demand.

Table 2: Descriptive Statistics

	Obs	Mean	Std. Dev.	Min	Max
Dependent Variable (y)					
Education spending	33507	3157.905	2094.114	2.607	52239.800
Controls (X)					
IDEB disclosure (D)	3723	0.7810	0.4136	0	1
gdp	33507	10202.330	9950.686	1320.591	342932.900
wage	33507	809.635	252.696	133.823	4173.603
tax price	33507	59.656	127.829	1.484	9574.092
categorical grants	33507	1558.334	665.505	0.035	7228.740
block grants	33507	597.468	443.957	0.065	6088.622
schooling	33507	9.744	1.254	2.500	15.828
occupation	33507	21.701	15.334	0.033	386.523
men	33507	50.618	1.485	44.170	67.744
population	33507	29.015	85.330	0.811	2505.554
elderly	33507	7.613	2.392	0.548	21.564
young	33507	35.732	6.467	16.616	62.556
rural	33507	31.207	30.167	0.000	100.000
second cycle	33507	24.232	19.483	0.000	100.000
competition	33507	3.160	1.574	1.000	20.000
incumbent's age	33507	48.989	9.623	20.000	89.000
incumbent's Education	33507	0.434	0.496	0.000	1.000
left	33507	0.214	0.410	0.000	1.000
incumbent women	33507	0.071	0.258	0.000	1.000
aldermen's education	33507	16.968	16.521	0.000	100.000
aldermen's age	33507	43.755	3.879	28.617	59.999
women in council	33507	12.097	10.898	0.000	77.778
competition for seats	33507	6.031	3.687	1.000	30.333
fragmentation	33507	74.584	9.675	0.000	94.230
majority of seats	33507	0.610	0.488	0.000	1.000
president's party	33507	0.076	0.265	0.000	1.000
governor's party	33507	0.221	0.415	0.000	1.000
lameduck	33507	0.305	0.460	0.000	1.000
Other					
n. of students x	3723	298.9699	852.7331	0	24770
More than 30 students (z)	3723	0.832	0.374	0.000	1.000

The demand for public services is also a function of income. Borcherding and Deacon (1972) and Bergstrom et al. (1982) find that public education is a normal good, i.e., it is increasing in income. However, for the period under analysis, there is no information about the mean income of the municipalities. Fortunately, however, some proxy variables are available, such as the GDP net of the public sector activities from IBGE, the average wage of formal sector workers from the Annual Relation of Social Information (RAIS) gathered by the Ministry of Labor and Employment, and an occupation index from the same source that consists of the ratio between the number of formal sector workers between 25 and 65 and the total number of individuals of the same age living in the municipality. Taken together, these three variables should capture the income effect on the demand for education. The average GDP per capita by municipality over the period is equal to 10,202 reals (5,859 USD) with significant dispersion. The average salary by municipality in the formal sector is 810 reals (465 USD), and the average percentage of people between 25 and 65 years old employed in the formal sector is only 21.7%. [20]

Other variables are included as controls to capture the differences in tastes for public education. For example, demographic variables such as the percentage of male, young and elderly individuals are related to tastes for public education. The proportion of men in the population is on average %50.62 and is included to account for the fact that men leave school for the labor market earlier than do women.[21] Elderly people (who amount to 7.61% of the total population), in turn, often demand less education and more health expenditures. The predominance of young people (who amount to 35.73% of the total population), however, can have an ambiguous effect on educational spending because it can lead to a higher demand for education (because localities will be populated by families with a strong preference for education expenditures) and at the same time it can diminish the spending over time if the cohorts become larger and the amount of resources per pupil becomes smaller.[22] In addition, years of schooling (9.74 years on average among the municipalities) is included to capture more-educated individuals' preference for more public education.

Local population (from IBGE) is considered to account for economies of scale in the provision of education. In principle, therefore, the larger the population is, the lower the per-pupil expenditure should be. However, large cities usually have higher costs of living that can affect expenditure levels. This phenomenon is difficult to address because there are no indexes that capture such peculiarities for all municipalities.

The percentage of students enrolled in rural public schools and the percentage of students enrolled in the second cycle of primary education (from 6th to 9th grade) at local public schools are included to take into account the differences in the amount of the categorical transfers that students attending urban schools receive (from FUNDEF) over the transfers provided to those enrolled in the first cycle (from 1st to 5th grade) or in preschool. The average proportion of rural students by municipalities is 31.2%. Of course, most of Brazil's municipalities have very small populations, with an important rural sector. Conversely, most

[20]Note that this proxy for occupation can be greater than the unity because the numerator refers to the total number of workers, whereas the denominator is restricted to citizens living in the municipality. Because many cities are predominantly residential, whereas the majority of jobs are concentrated in few cities, it is expected that few cities present indicators greater than 1, whereas the majority present indicators smaller than 1. Because we are averaging the indexes over the municipalities and because the small municipalities, which comprise the majority of municipalities, have a small proportion of formal jobs, the occupation rate in the formal sector must be underestimated.

[21]A report by the OECD (2009) shows that the difference in the upper secondary graduation rates for boys and girls of the appropriate age is especially remarkable in Brazil, at 71,9% among girls and 53,2% among boys.

[22]Poterba (1997) briefly discuss this subject.

of the population is concentrated in a few municipalities, which are predominantly urban. Thus, the proportion of rural students in Brazil is well below one-third of the population.

The political variables included in the model are considered by the literature as important determinants of the level of expenditures. Left-wing governments, for example, prefer a larger public sector, i.e., higher expenditures[23], although in Brazil, partisan ideology is not as well defined as in other countries.[24] Nevertheless, a dummy variable named "left" is included in the model to capture the possible differences in tastes for public expenditures on education. In addition, two dummy variables assigning the incumbent's party alignment with presidential or gubernatorial parties account for the fact that the incumbents' partisanship is supposed to increase the amount of resources to which they have access.

We also include a variable of party fragmentation (see the description in Table 1). A more fragmented political system supposedly reflects the existence of various interest groups. According to Weingast et al. (1981), because resources come from a common pool of taxation, any expenditure that is targeted at specific groups will have its costs equally divided among all groups, causing the costs of the program not to be fully internalized by the benefited groups and thus increasing the demand for public spending. Additionally, the incumbents can engage in pork-barrel politics to overcome the difficulties imposed by a fragmented city council, thus increasing spending levels.

A dummy variable indicating the term of the incumbent (i.e., lame duck or not) is also considered in the empirical model. Besley and Case (1995) argue that lame-duck incumbents have an adverse incentive to maximize rent extraction because they need not run for reelections, which translates into higher taxes and expenditures during the final term in office.

Mukherjee (2003) estimates that the size of the majority can also affect the level of public spending.[25] Our strategy considers a dummy variable, assigning the value of 1 when the party in office holds more than 50% of the city council's seats (and 0 otherwise). This dummy variable controls for the average effect of the political majority on educational spending. A set of incumbent and aldermen's characteristics that are intended to reflect their quality

[23]See Alt and Lowry (1994) and Sollé Ollé (2006).

[24]See Lucas and Samuels (2010).

[25]According to that author, weak majorities (greater than 50% and smaller than 56%) lessen the need to engage in pork-barrel politics and thus decrease the level of total expenditures. Conversely, strong majorities (between 56% and 68%) can diminish the risks of adopting loose fiscal policies and transfer the burden to non-majority members. However, when a supermajority (greater than 68%) is achieved, the burden cannot be passed on to the minority group because it is too small and increasing expenditures with a budget restriction means that the majority will have to cope with the costs of taxation.

is also included in the econometric model. One of these characteristics is the education of incumbents and aldermen, which can reflect their preferences regarding educational expenditures. Besley and Case (1995) also emphasize incumbents' age as an important determinant of electoral outcomes and fiscal policy. Incumbents on the verge of retirement, who may be serving a final term, have an incentive to extract more rents and thus to increase both taxes and spending levels. Therefore, both incumbents' age and aldermen's mean age are used as additional controls.

Milyo and Schosberg (2000) demonstrate that because women face barriers to entering office, if they are chosen, they can be claimed to be of better quality. Therefore, we build two variables to capture this phenomenon: a dummy variable indicating whether the mayor is a woman and the percentage of women on the city council.

Finally, competition can also lead to higher-quality incumbents. As the number of candidates for a position increases, voters are better able to distinguish between good and bad candidates. Consequently, the expected rent extraction is smaller, but we cannot unequivocally determine that better incumbents will tax and spend less. Accordingly, one variable that informs the number of candidates running for office and another that reflects the number of candidates per council seat are built.

4 Preliminary Data Analysis

This section brings the results of a preliminary analysis of the panel data and relies on a spatial autoregressive model with an additional interaction term between the spatial lag and a dummy for the post-IDEB disclosure period. The model below also represents the demand for education[26]. Assume

$$\begin{aligned} y_{it} = \alpha + \lambda_0 \sum_{j=1}^{N} W_{ij} y_{jt} + \lambda_1 T_{0813} \sum_{j=1}^{N} W_{ij} y_{jt} + \\ X_{it}\beta + \mu_i + \tau_t + u_{it}, \end{aligned} \tag{1}$$

where y_{it} denotes educational spending per pupil. Element D represents a dummy that equals 1 from 2008 to 2011 and 0 otherwise (between 2003 and 2007). The W_{ij} term is

[26]See Borcherding and Deacon (1972) and Revelli (2006).

the spatial weight assigned to unit j by the unit i defined by the contiguity criterion. The weights result from the row standardization of the $N \times N$ spatial weights matrix W_N such that $\sum_{j=1}^{N} W_{ij} = 1$ for the $i-th$ row. The neighbors' educational expenditure per pupil is represented by $\sum_{j=1}^{N} W_{ij} y_{jt}$. The coefficient λ_0 informs the spatial correlation before the IDEB disclosure and λ_1 informs the difference between the spatial correlation before and after the results of the index were made public, with $|\lambda_0|, |\lambda_1| < 1$ to ensure spatial stationarity. Vector X_{it} is $1 \times K$ and represents the demographic and political covariates, whereas β is a $K \times 1$ vector of corresponding parameters.

Element μ_i represents the spatial-specific effects and is aimed at capturing the non-observable characteristics that do not vary over time but that are potentially correlated with the covariates in the model. A spatial Hausman test is performed to decide whether μ_i is fixed or random. Common shocks to all municipalities at a given point in time are represented by τ_t, a set of year dummies, and the random component of the composite error is given by u_{it}.

We can estimate the spatial lag model in 1 either by maximum likelihood or using the instrumental variable approach. There are pros and cons associated with each method. The primary advantage of the former method consists of being efficient and restricting the spatial parameter to lying between -1 and 1. Conversely, if we have non-spherical disturbances, the variance estimator of maximum likelihood coefficient will produce inconsistent estimates, unless we can model the disturbance autocorrelation and heteroskedasticity.

If we estimate the model by maximum likelihood, we assume the error term is such that $u \sim \mathcal{N}(0, \sigma_u^2)$.[27] The Log Likelihood function that provides the estimates of the spatial parameters and the other coefficients is given by

$$\begin{aligned} lnL = &-\frac{NT}{2} ln(2\pi\sigma^2) + \sum_{t=1}^{T} ln|I_N - \lambda_0 W_N - \lambda_1 D W_N| \\ &- 1/(2\sigma^2) \sum_{i=1}^{N}\sum_{t=1}^{T} \left[y_{it} - \lambda_0 \sum_{j=1}^{N} W_{ij} y_{jt} - \lambda_1 D \sum_{j=1}^{N} W_{ij} y_{jt} \right. \\ &\left. -\alpha - X_{it}\beta - \mu_i - \tau_t \right]^2 \end{aligned} \tag{2}$$

[27]In fact, even if we reject the null hypothesis that the disturbances are normally distributed Lee (2004) shows that the parameters can be asymptotic and normally distributed under weak regularity conditions.

where D is a $N \times N$ diagonal matrix whose diagonal elements are the regime dummies D_{it} for each cross-sectional unit at time t. The parameters are then estimated by maximizing the profile likelihood function concentrated with respect to the parameters of the exogenous variables and the variance of the disturbance.

Several authors assume that voters react to differences in fiscal policies due to observable characteristics because they are not perfectly informed.[28] In this case, the spatial lag model would be preferable. Conversely, Bordignon et al. (2003) argue that the spatial correlation in the error term makes more sense than spatial correlation in the dependent variable. The authors reason that voters have enough information to avoid being influenced by differences between jurisdictions' tax and expenditure levels due to observable characteristics. Instead, they are more likely to evaluate neighbors' unexpected changes in public policies.[29] In this case, the appropriate model to identify yardstick competition would be one of the spatial error type. Ultimately, though, the choice of model type can be considered an empirical issue that is made based on the robust LM lag and the LM error tests proposed by Anselin et al. (1996).

If we confront the spatial-lag and spatial-error models, the former has the advantage of producing consistent estimates of the coefficients even when it is not the correct model. If there is autocorrelation in the residuals, the coefficients' standard errors will be inconsistent. One safe alternative to obtain valid coefficients and standard errors is to estimate the spatial lag model by the generalized method of moments with a heteroskedasticity- and autocorrelation-consistent covariance matrix (GMM-HAC). Note, however, that unlike the maximum likelihood model, the instrumental variable approach does not restrict the spatial parameter to lying between -1 and 1.

To represent the GMM model, let the spatial matrix be such that $W = I_T \otimes W_N$, and $H = [1,\ Wy,\ DWy,\ X,\ \tau]$ be the $NT \times L$ matrix of regressors, $\delta = [\alpha,\ \lambda_0,\ \lambda_1,\ \beta,\ \kappa]'$ be the $1 \times L$ vector of parameters, and let $Z = [1,\ WX,\ DWX,\ X,\ \tau]$ be the $NT \times M$ matrix of instrumental variables for H (with $M > L$), including first-order spatial lags of the independent variables, their interactions with the regime dummy D, the matrix X, that serves as the instrument for itself, and year dummies given by the vector τ. Model 1 can be rewritten as

$$Y = H\delta + \mu + u \tag{3}$$

[28] See Revelli (2006), Elhorst and Fréret (2009), Allers and Elhorst (2005), Sollé Ollé (2003), and Revelli (2002).

[29] See Besley and Case (1995) and Revelli and Tovmo (2007).

where μ represents the specific part of the composite error (i.e., the fixed effect), and u represents the random error. GMM estimation involves minimizing a quadratic function of the moment conditions

$$Q(\delta) = \overline{m}(\delta)'[Var(\overline{m}(\delta))]^{-1}\overline{m}(\delta) \tag{4}$$

where the moments $\overline{m}$ of the demeaned variables (represented by two dots over the variables) are given by

$$\overline{m}(\delta|\ddot{Y}, \ddot{H}, \ddot{Z}) = \frac{1}{NT}\ddot{Z}'\ddot{u} = \frac{1}{NT}\ddot{Z}'(\ddot{Y} - \ddot{H}\delta) \tag{5}$$

and the variance-covariance matrix $Var(\overline{m}(\delta))$ allows autocorrelation and heteroskedasticity.

4.1 A simple spatial model ($\lambda_1 = 0$)

Table 3 contrasts spatial and non-spatial models. The first two columns in Table 3 present models in which the spatial parameters (in the error or in the dependent variable) are all set to zero. The first model (POLS) is estimated using ordinary least squares and is only illustrative of the importance of considering the fixed effects of the municipalities. As observed in the bottom of Table 3, the robust LM lag and the LM error test statistics performed with the residuals of the models reject neither a spatial lag nor a spatial error model as the best suited for the problem. The Moran's I calculated on the residuals indicates a spatial auto-correlation equal to 0.2831, which is significant at less than the minimum conventional level of 1%.

As noted in Elhorst (2010), failing to take fixed effects into account can result in spatially auto-correlated residuals. The fixed-effect model (Within) in the second column clearly confirms that finding. The Moran's I that is calculated on the residuals of the fixed effects model is estimated as 0.1639, which remains significant at less than 1% but is considerably smaller than the correlation observed in the residuals of the POLS model. We also obtain smaller statistics on both robust LM lag and LM error tests, reflecting the smaller spatial correlation after removing the fixed effect. In any event, spatial auto-correlation remains, and we reject the hypothesis of no spatial lag at a higher significance level than we reject the null of no spatial error correlation, thus indicating the spatial lag as the most appropriate model.

Next, we estimate a fixed-effects model with a spatial lag by maximum likelihood. The

Hausman test at the bottom of Table 3 indicates that the fixed-effects estimates are different from those of random effects, thus favoring the choice of the consistent estimator. Compared to the other models in Table 3 the ML model presents the higher log-likelihood value. Therefore, we could say that for normally distributed disturbances, this would be the best model to fit the data. The estimated spatial correlation coefficient of 0.287 indicates a substantial amount of interaction in educational spending, a figure similar to that found by Revelli (2006) (estimated as 0.216) for welfare spending and much higher than that found by Elhorst and Fréret (2009) (equal to 0.083). However, notwithstanding the fact that the log-likelihood statistics and the robust LM lag and LM error tests suggest that the maximum likelihood spatial lag model would best describe the true model, we cannot ignore the spatial correlation in the error term. Ignoring that does not bias the coefficients but instead produces inconsistent standard errors.

Table 3: Non-Spatial vs. Spatial models

	POLS	WITHIN	SARFE	
			ML	GMM-HAC
Wy			0.287***	0.791***
			(45.397)	(34.360)
gdp	0.260***	0.100***	0.087***	0.052***
	(50.420)	(9.045)	(11.421)	(5.631)
wage	0.356***	0.179***	0.141***	0.071***
	(33.262)	(13.137)	(16.032)	(7.378)
tax price	-0.075***	-0.060***	-0.055***	-0.034***
	(-20.157)	(-6.702)	(-12.895)	(-5.267)
categorical grants	0.183***	0.107***	0.100***	0.047***
	(11.109)	(5.869)	(32.008)	(3.340)
block grants	0.192***	0.040*	0.046***	0.032
	(15.182)	(1.798)	(5.080)	(1.523)
schooling	0.042***	0.023	0.003	-0.025
	(3.235)	(0.906)	(0.212)	(-1.443)
occupation	0.009***	0.005	0.004	0.001
	(2.623)	(1.356)	(1.605)	(0.291)
men	0.001	-0.016***	-0.012***	-0.005**
	(1.052)	(-5.396)	(-6.454)	(-2.260)
population	-0.003	0.090***	0.080***	0.027
	(-0.489)	(3.261)	(4.860)	(1.327)
elderly	-0.012***	-0.057***	-0.042***	-0.015***
	(-9.474)	(-17.756)	(-24.404)	(-6.931)
young	-0.036***	-0.034***	-0.025***	-0.008***
	(-56.255)	(-19.121)	(-22.782)	(-6.089)
rural	0.002***	0.001***	0.001***	0.001***
	(30.398)	(6.284)	(14.567)	(9.090)
second cycle	-0.004***	-0.001***	-0.001***	-0.001***
	(-45.514)	(-2.907)	(-6.756)	(-7.025)
competition	0.017***	0.077***	0.053***	0.013***
	(3.916)	(16.136)	(15.754)	(3.385)
incumbent's age	-0.028***	-0.005	-0.001	0.004
	(-3.786)	(-0.524)	(-0.101)	(0.529)
incumbent's Education	-0.005*	0.007	0.006**	0.004
	(-1.763)	(1.592)	(2.022)	(1.394)
left	0.023***	0.021***	0.018***	0.015***

Continued on next page

Table 3 – *Continued from previous page*

	POLS	WITHIN	SARFE	
			ML	GMM-HAC
	(5.366)	(3.515)	(5.030)	(3.801)
incumbent women	-0.030***	-0.011	-0.009*	-0.011**
	(-6.060)	(-1.405)	(-1.786)	(-2.179)
aldermen's education	0.001***	0.000**	0.0003**	0.000
	(12.424)	(2.352)	(2.395)	(0.525)
aldermen's age	-0.169***	-0.059**	-0.041**	-0.008
	(-8.838)	(-2.359)	(-2.304)	(-0.419)
women in council	-0.001***	0.000	0.0000	-0.000
	(-6.122)	(0.107)	(-0.084)	(-0.969)
competition for seats	0.016***	-0.021**	-0.016***	-0.005
	(3.430)	(-2.561)	(-2.811)	(-0.960)
fragmentation	-0.001***	-0.000	-0.0003*	-0.000
	(-4.737)	(-0.930)	(-1.807)	(-0.811)
majority of seats	-0.001	0.008**	0.006**	0.005*
	(-0.231)	(2.340)	(2.326)	(1.846)
president's party	0.032***	-0.000	-0.001	-0.005
	(4.587)	(-0.013)	(-0.158)	(-0.775)
governor's party	0.002	0.002	0.002	0.003
	(0.653)	(0.430)	(0.816)	(1.275)
lameduck	0.008**	0.009***	0.008***	0.006**
	(2.415)	(3.043)	(3.421)	(2.437)
Spatial Fixed Effects	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Excluded Instruments	-	-	-	WX
Robust LM no Spatial Lag χ^2	34668.448***	1178.3453***	-	-
Robust LM no Spatial Error χ^2	39285.841***	418.5705***	-	-
Spatial Hausman χ^2	-	-	1252.113***	
N	33507	33507	33507	33507
log-likelihood	-3496.3444	14543.129	15595.582	15076.04

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
Dependent variable y is the education spending. The spatially lagged dependent variable Wy uses the contiguity criteria to assign neighborhood.
The endogenous variable is the spatially lagged dependent variable Wy.
The instruments are the first order spatial lags of the regressors in WX.
Full estimates can be obtained upon request to the authors.

One way to overcome this problem of spatially correlated residuals is to estimate the spatial lag model with robust standard errors, which can be implemented with the generalized method of moments with heteroskedasticity and autocorrelation robust standard errors (GMM-HAC).[30] The instrument consists of the first order spatial lag of the regressors. [31] The GMM-HAC model in Table 3 shows a spatial parameter of 0.791, which is considerably higher than the previously mentioned results, most likely because of the linear relationship between the endogenous regressor and the dependent variable that is peculiar to IV estimators.

[30]Contrary to the maximum likelihood coefficients, the IV coefficients of the exogenous variables can be interpreted as marginal effects.

[31]We use only the first-order spatial lag of the control variables as instruments so that we can compare our models in this with those in the next sections. Note also that when we run models with observations within a small bandwidth, we do not have enough degrees of freedom to use second-order neighbors as instruments.

The coefficients on variables gdp, wage and occupation are all positive in the GMM-HAC model. Because these variables serve as a proxy for income, we expect a positive sign. The magnitude of the elasticity of GDP with respect to educational spending is quite small (0.052). The same is true for the wage elasticity (0.071).

Surprisingly, the estimated elasticity of the categorical grants with respect to educational spending is 0.047. That means that a marginal dollar increase in the earmarked transfers for education will be offset by a reduction of almost one dollar in the general-purposes resources that have been put into education. One possible explanation for this phenomenon is that the income elasticity of the demand for public education is low and/or that local governments have already spent at optimal levels and therefore, increasing the amount of money available will not lead an additional increase in public spending. The elasticity of the block grants is even smaller (0.032) and non-significant, which reinforces this conclusion, i.e., an increase in grants that are not earmarked for education will not increase the spending level on education for this period and sample.

The tax price shows a negative effect on educational spending per pupil, meaning that the higher the cost perceived by the citizens of spending an additional monetary unit on public education, the smaller the demand for this good. Schooling does not appear to have a significant effect once the fixed effects are controlled for. The same is true for the coefficient on population. A higher proportion of men in the municipality appears to reduce educational spending, possibly because they tend to leave school early to work and do not value education as much as women do.

Both the percentages of the elderly and the young present negative coefficients. The first result is direct: elderly citizens demand less education and more health expenditures.[32] However, the second result has a less obvious interpretation. One would expect a larger fraction of young people in a jurisdiction to increase the demand for education. However, if municipalities with many young individuals raise little revenue — due the lower share of economically active individuals —, an increase in the share of young individuals could mean that there will be fewer resources per capita available to finance education.

The coefficient on the percentage of students of the local public educational system enrolled in rural schools is positive, reflecting the legal determination that the rural school students receive a greater amount of transfers from FUNDEF. Conversely, the coefficient on the percentage of public school students enrolled in the 2^{nd} cycle (6th to 9th grade of fundamental education) is negative, despite the fact that the law requires higher transfers to

[32] See Poterba (1997) and Arvate and Zoghbi (2010).

those students. This could mean that the earlier stages of education demand more complementary expenditures than do the later stages, or people prefer spending on earlier stages of education.

The characteristics of the incumbents do not seem to be important determinants of education expenditures. For example, the age and education coefficients are not statistically significant.[33] Gender, however, appears to have a small effect. Women incumbents spend 1.1% less on education, according to our estimates.

Leftist incumbents spend 1.5% more than rightist incumbents and an increase of one candidate in the competition for the mayor's seat raises the education expenditure by 1.3%. Lame-duck incumbents spend 0.6% more on education and incumbents holding a majority of seats on the council spend 0.5% more on this expenditure function. Party fragmentation, competition for seats in the legislature, and the alignment between the mayoral party and the gubernatorial or presidential party have no statistically significant coefficients. Finally, aldermen's personal characteristics have little effect on educational spending, e.g., age, education and gender do not significantly affect educational spending.

4.2 Post-IDEB strategic behavior ($\lambda_1 \neq 0$)

This subsection verifies whether the IDEB disclosure affected strategic behavior between neighboring governments. It consists of a before and after exercise as in Revelli (2006). We expect that spatial correlation in educational spending resulting from yardstick competition to be reduced after municipalities' education indexes and their goals were made public. This is because the IDEB and its goals consisted of providing voters with new information about what is a high-quality education. This new information must have enabled voters to determine whether their mayors were good without having to look at what was happening to their neighbors.

Table 4 show maximum likelihood and GMM-HAC estimates. The signs of the estimated coefficients are identical, but the magnitude differs a bit. The spatial correlation is high in the ML model and even higher in the GMM-HAC model. Interestingly, both models suggest that the spatial correlation was reduced after the disclosure of the indexes — captured by the coefficient on the interaction between the dummy of period D and the spatial lag.[34] The

[33] Actually, in the maximum likelihood model, the incumbent's education has a positive influence on the expenditure level, but the standard errors of this model may not be reliable.

[34] We consider the year of 2008 as the first year post-IDEB disclosure. Remember that the indexes were

ML estimator suggest that the reduction was equal to 0.075 correlation point. The GMM-HAC estimator, however, shows an even higher reduction in spatial autocorrelation after the disclosure of the IDEB. According to that estimator, spatial correlation dropped 0.1242 points after the indexes became public information. With more information about education quality such as the national ranking provided by municipalities' IDEBs and the yearly goals defined by the Ministry of Education, the distribution of information about education quality between voters and incumbents became less asymmetric. Our interpretation is that voters stopped relying so much in their neighbors' educational spending as benchmarks of educational quality. This provides a possible explanation for the reduction observed in the spatial correlation in educational spending.

Table 4: Spatial interaction in the Post IDEB disclosure period

	Post IDEB disclosure	
	ML	GMM-HAC
Wy	0.332***	0.6166***
	(37.174)	(20.5862)
$D \times Wy$	-0.075***	-0.1242***
	(-6.207)	(-10.9156)
Controls	Yes	Yes
Spatial Fixed Effects	Yes	Yes
Time Fixed Effects	Yes	Yes
N	33507	33507
log-likelihood	15614.768	16226.001

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
Dependent variable y is the education spending. The spatially lagged dependent variable Wy uses the contiguity criteria to assign neighborhood.
Endogenous variables are the spatially lagged dependent variable Wy and the interaction DWy. The control variables are: gdp, wage, tax price, categorical grants, block grants, schooling, occupation rate, % men, population, % elderly, % young, % rural, % second cycle, competition, incumbent's age, incumbent's Education, left, incumbent women, aldermen's education, aldermen's age, women in council, competition for seats, fragmentation, majority of seats, president's party, governor's party, lameduck.
The instruments are the first order spatial lags of the regressors and its interaction with the disclosure variable $[WX,\ DWX]$. Full estimates can be obtained upon request to the authors.

The results presented in this section agree with those of Revelli (2006) for welfare expenditures, i.e., that strategic interactions in educational spending have decreased after the broad disclosure of the indexes that improved the available information about educational quality and diminished the importance of local information spillovers in voters' decisions. However, Revelli (ibid) recognizes that the empirical evidence found in his work reflects a situation at a given point in time in the sense that he was able to build a panel with only one period immediately before and another immediately after the introduction of a national performance rating of social expenditures. Other non observed factors occurring concomitantly with given institutional changes may have played some role.

disclosed in mid-2007, but the education budgets were already defined for that year when the indexes were unveiled. Thus, the publication of the index must have produced effects from 2008 on.

4.3 The political nature of the heterogeneity in the spatial correlation coefficient

We still cannot tell whether the spatial correlation is due to yardstick competition, welfare competition or other competing theories. Yardstick competition arises not only out of information asymmetry between voters and incumbents but also out of incumbents' political incentives. Welfare competition models arise out of ruling officers' efficiency concerns. In Table 5, we present a test of the nature of the spatial process. We interact the spatial parameters with dummies of lame-duck incumbents and mayors with majority support in the legislature. Each of these binary variables represents distinct political incentives for mimicking neighbors' educational spending. Both maximum likelihood models and GMM-HAC present similar results. The main difference lies in the magnitude of the spatial parameters estimated by each method.

The results provide evidence of yardstick competition in educational spending. Municipalities with Lame-duck incumbents show smaller spatial correlation, so they tend to interact less with their neighbors in a strategic manner. This result is typical of yardstick competition. The incentive for an incumbent in his last term to mimic a neighbor to signal his quality to voters is weaker.This result is consistent with those of Besley and Case (1995) and Bordignon et al. (2003), who analyze interactions in tax-setting and predict fewer incentives for lame-duck incumbents to mimic their neighbors' behavior and signal their quality to the voters. Note also that an incumbent being a lame duck does not eliminate spatial interaction, thus supporting the argument of Alesina and Spear (1988), according to which the parties have incentive mechanisms to prevent lame-duck governors from pursuing only their own interests. Likewise, when incumbents hold the majority in the legislature, they tend to act less strategically and are less likely to imitate their neighbors. Incumbents without majority support will find it advantageous to imitate their neighbors because of an uncomfortable political situation that requires them not to get behind their neighbors — who serve as benchmark to the voters.[35] [36]

[35]Elhorst and Fréret (2009) also find similar results for welfare spending but instead, they consider majority governments to be those whose incumbents have the support of more than 75% of the aldermen.

[36]Note that the difference in the magnitude of the spatial parameters estimated by Maximum Likelihood or Instrumental Variables must be due to the non-linear nature of the spatial lag.

Table 5: Heterogeneity of the spatial parameter and the political incentives to engage in Yardstick Competition

	Lame Duck		Majority Support Legislature	
	ML	GMM-HAC	ML	GMM-HAC
Wy	0.320***	0.797***	0.329***	0.790***
	(24.981)	(35.635)	(37.721)	(34.577)
Lame Duck $\times$ Wy	-0.045***	-0.013***		
	(-2.775)	(-2.962)		
More than 50% $\times$ Wy			-0.104***	-0.012**
			(-6.629)	(2.357)
Controls	Yes	Yes	Yes	Yes
Spatial Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
N	33507	33507	33507	33507
log-likelihood	15597.963	14955.697	15619.869	15073.435

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
Dependent variable y is the education spending. The spatially lagged dependent variable Wy uses the contiguity criteria to assign neighborhood.
Endogenous variables are the spatially lagged dependent variable Wy and the interactions Lame Duck $\times$ Wy and More than 50% $\times$ Wy.
The control variables are: gdp, wage, tax price, categorical grants, block grants, schooling, occupation rate, % men, population, % elderly, % young, % rural, % second cycle, competition, incumbent's age, incumbent's Education, left, incumbent women, aldermen's education, aldermen's age, women in council, competition for seats, fragmentation, majority of seats, president's party, governor's party, lameduck.
The instruments are the first order spatial lags of the regressors and its interaction with the variables Lame Duck or More than 50% ([Lame Duck $\times$ WX,More than 50% $\times$ WX]).
Full estimates can be obtained upon request to the authors.

5 Identification Strategy: The IDEB's Disclosure Rule

The preliminary analysis cannot rule out the possibility that some unknown confounding factor may have occurred concomitantly with the IDEB disclosure, which would render biased the differences in the spatial parameter before and after the IDEB. Moreover, because participating in both the Plan of Goals "All Committed to Education" and the standardized test of *Prova Brasil* — which is used to calculate IDEB — is voluntary, some municipalities can choose whether to have their IDEB disclosed.Thus, one could argue that the decision to participate and to have the index published is endogenous in the model.

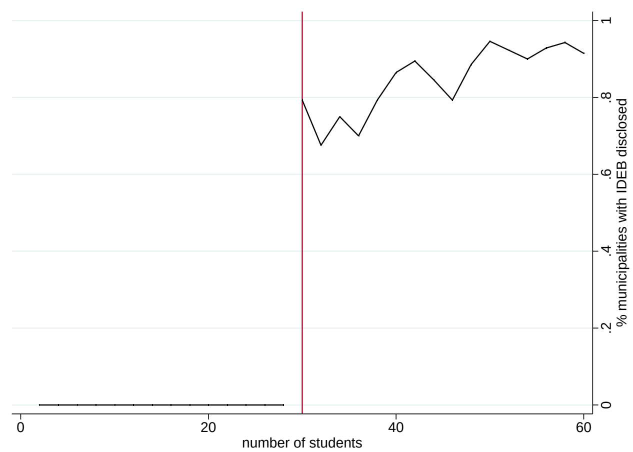
Fortunately, however, to disclose the first wave of the index — in 2007 —[37], the Ministry of Education established a cutoff of 30 students enrolled and present on the day of *Prova Brasil's* exam, to minimize the sampling error associated with the schools' average performance. [38] This cutoff of 30 students provides an exogenous jump in the probability

[37]The first time the IDEB and its goals were disclosed, in April 2007, it was calculated with data from the first edition of *Prova Brasil* of November 2005 and the Education Census. Because the 2007 education budget was already defined when IDEB was first disclosed, we use 2008 cross-section data on educational spending.

[38]Four hundred and ninety-six municipalities — out of 5,564 — with more than 30 students enrolled

of participating in the exam and consists of a potentially good instrument for the IDEB's disclosure dummy. Within a sufficiently small bandwidth around this cutoff, both observable and non-observable characteristics are independent of the treatment status, as if the treatment (IDEB's disclosure) had been randomly assigned.

Subsequent waves of IDEB and its goals are less useful in this case because later on, the Ministry of Education changed the disclosure rule to 20 students, significantly diminishing the sample size around the cutoff. Moreover, using only the first wave of data on IDEB reduces the possibility of manipulation of the number of students (forcing variable). In the first wave of IDEB — disclosed in mid-2007 — education accountability was still a novelty. In the following years, knowing the pros and cons of participating in the *Prova Brasil* exam, it is possible that some municipalities have manipulated the number of students to participate (or not) in the assessment.



Source: Elaborated by the author using data from the Ministry of Education.
Figure 2: Proportion of municipalities with IDEB disclosed according to the number of students

To consistently estimate the effects of disclosing the IDEB on the strategic interaction in educational spending we estimate the following local regressions in two-stages:

in the municipal school system did not take part in the first wave of the assessment. Nine hundred and forty-two municipalities had no students enrolled in the municipal school system and were excluded from the initiative. Thus, 1,438 municipalities did not participate in the first wave of the assessment of the IDEB and its goals.

Second Stage

$$y_i = \alpha_0 + \lambda_0 W_i y + \lambda_1 D_i W_i y + \gamma_0 D_i + \theta_0 D_i (x_i - c) + \phi_0 (x_i - c) + X_i B_0 + u_i \tag{6}$$

First Stage

$$\begin{aligned}
W_i y &= \alpha_1 + W_i X \delta_1 + z_i W_i X \delta_2 + \gamma_1 z_i + \theta_1 z_i (x_i - c) + \phi_1 (x_i - c) + X_i B_1 + \nu_i \\
D_i &= \alpha_2 + W_i X \delta_3 + z_i W_i X \delta_4 + \gamma_2 z_i + \theta_2 z_i (x_i - c) + \phi_2 (x_i - c) + X_i B_2 + \eta_i \\
D_i W_i y &= \alpha_3 + W_i X \delta_5 + z_i W_i X \delta_6 + \gamma_3 z_i + \theta_3 z_i (x_i - c) + \phi_3 (x_i - c) + X_i B_3 + \epsilon_i \\
D_i (x_i - c) &= \alpha_4 + W_i X \delta_7 + z_i W_i X \delta_8 + \gamma_4 z_i + \theta_4 z_i (x_i - c) + \phi_4 (x_i - c) + X_i B_4 + \varepsilon_i \\
s_i \in S &= \{c - h < s_i < c + h\} \quad h =, 5, 6, 7, 10, 20 \text{ and } 30
\end{aligned} \tag{7}$$

where W_i is a $1 \times N$ vector of the i-th municipality's neighbors, y is the $N \times 1$ vector of regressand observations, X is the $n \times K$ matrix of regressors, B_k is the vector of parameters and D_i is the dummy of IDEB publication.[39] The first-stage regressions generate exogenous fitted variables to enter the second-stage regressions that provide the parameters of interest. Because $W_i y$ and D_i are endogenous (the former by construction and the latter is a possibility, but we are conservatives) and so the interaction terms $D_i W_i y$ and $D_i(x_i - c)$, we need exogenous instruments to estimate the parameters consistently. The neighbors' K covariates in $W_i X$ are the natural candidates for instrumenting $W_i y$. The instrument for the IDEB disclosure dummy D_i is represented by z_i, which assigns a value of 1 for school systems with more than 30 students enrolled and 0 otherwise. The instruments for the interaction term $D_i W_i y$ are given by the vector $z_i W_i X$, and the instrument for $D_i(x_i - c)$ is $z_i(x_i - c)$. Finally, the idea is to estimate this regression for municipalities with a number of enrollments $s_i \in S$, i.e., within bandwidths $h = 5, 6, 7, 10, 20$ and 30 students around the cutoff $c = 30$ students. Such a procedure allows the evaluation of similar municipalities on each side of the cutoff, minimizing differences in terms of unobservable characteristics correlated with the disclosure (treatment) variable.

[39] There are two types of IDEB for each school and municipality, one calculated with the average proficiency of the $5th$ graders, which is called "First cycle IDEB", and the other calculated with the average proficiency of the $9th$ graders, which is called "Second cycle IDEB". We focus our analysis on the disclosure of the "First cycle IDEB" because in most municipalities, the "Second cycle" grades — $6th$ to $9th$ grades — remain under the administration of the states instead of the municipalities.

Table 6 reflects the remarkable shrinkage of differences in the characteristics of treatment and untreated municipalities when one restricts the observations to lie in a small interval — of the forcing variable — around the cutoff. We evaluate 6 intervals, $h = 5, 6, 7, 10, 20, 30$. In fact, when $h = 10$ much of the heterogeneity in the mean characteristics of municipalities in both sides of the cutoff vanishes. In the smallest interval $h = 5$, only the age of the incumbent and the proportion of women in council remain different in both sides of the cutoff. This means that much of the unobserved factors must also be similar in both sides the cutoff, reducing the possibility of inconsistent estimates regarding the disclosure of IDEB.

Table 6: Mean difference between covariates of municipalities to the right and to the left of the cut-off ($h = 5, 6, 7, 10, 20, 30$) in 2008

	h=5	h=6	h=7	h=10	h=20	h=30	all
gdp	0.052	0.022	0.054	0.006	-0.054	-0.06	-0.104***
wage	0.012	0.006	0.003	-0.02	-0.02	-0.027	-0.018*
tax price	0.082	0.06	0.069	0.069	-0.001	0.028	0.57***
categorical grants	0.021	0.027	0.035	0.021	-0.013	-0.024	-0.066***
block grants	-0.01	-0.03	-0.021	-0.064	-0.066	-0.094**	-0.631***
schooling	0.031	0.026	0.022	0.004	-0.002	-0.002	-0.007
occupation	0.076	0.056	0.102	0.081	0.056	0.063	0.256***
men	-0.36	-0.584**	-0.648***	-0.444**	-0.318**	-0.247*	-0.917***
population	-0.011	0.034	0.028	0.088	0.088	0.119**	1.132***
elderly	0.313	0.406	0.58*	0.26	-0.275	-0.539***	-1.804***
young	0.122	0.203	-0.329	0.18	1.047**	1.456***	2.787***
rural	-5.204	-5.312	-5.598	-4.918	-8.395***	-11.654***	-35.853***
second cycle	0.754	-0.744	0.701	-0.007	-1.02	-0.632	1.23
competition	0.074	0.056	0.054	0.056	0.049**	0.051**	0.201***
incumbent's age	-0.065**	-0.05*	-0.042	-0.041*	-0.037**	-0.033**	0.019**
incumbent's Education	0.054	0.048	0.04	-0.012	-0.003	0.014	0.13***
left	0.013	0.036	0.013	0.037	0.037	0.013	0.03*
incumbent women	0.055	0.045	0.031	0.038	0.007	0.024	0.022**
aldermen's education	0.174	1.084	1.329	1.239	1.568	1.481	9.477***
aldermen's age	0.008	0.006	0.012	0.01	0.003	0	0.018***
women in council	-3.713**	-3.06**	-2.941**	-1.22	-0.57	-0.059	0.588
competition for seats	0.085	0.084	0.066	0.11**	0.109***	0.105***	0.501***
fragmentation	1.602	1.346	0.138	1.08	1.609*	1.882**	4.813***
majority of seats	0.03	0	0.038	0.024	0.009	0.03	-0.114***
president's party	0.052	0.054	0.039	0.043	0.03	0.01	0.016
governor's party	0.043	0.001	0.019	0.004	0.063*	0.069**	0.046***
lameduck	-0.059	-0.067	-0.038	-0.048	-0.041	-0.013	-0.028
N	173	205	230	312	582	799	3723

Notes: *** Difference statistically significant at 1%; ** Difference statistically signifficant at 5%; * Difference statistically significant at 10%.

6 Main Results

This section explores the discontinuity in the enrollment-based rule that determines the unveiling of the IDEB (in mid-2007) for municipalities with more than 30 students. First, we present the Regression Discontinuity Design results without taking into account the possibility of spatial interactions between municipalities. That is, Table 6 results only assess whether the local governments increase or decrease their educational expenditures in response to the IDEB's disclosure within bandwidths $h = 5, 6, 7, 10, 20 \quad \text{and} \quad 30$.

Table 7 shows the first stage estimates for the disclosure variable D of the models in Table 8. The exogenous instrument for D is z, the dummy assigning 1 to municipalities with more than 30 students enrolled in the grade under assessment. The instrument for $D \times (x - c)$ ist $z \times (x - c)$, that allows different slopes on each side of the cutoff. [40] These first stage estimates reveal a very strong jump — about 70% — in the probability of having IDEB disclosed for those to the right of the cutoff. Thus we have a strong instrument around the cutoff.

The coefficients on variable D in Table 8 inform the treatment effect. Starting from the the greater bandwidth of $h = 30$ the education spending increased 7.3% in response to the disclosure of IDEB, but in model (12), with covariates. Restricting the bandwidth to $h = 20$ actually improves the significance of the treatment effect. Model (9), without covariates, shows the disclosure have an impact of 21.8% on education spending, whereas model (10) suggests a smaller effect, of 12%. If we narrow the bandwidth further, to $h = 10$, the estimated impact of IDEB's disclosure on the education expenditure is only marginally significant, in model (7) the estimates reaches 25% without covariates while it remains at 12% in model (8), with covariates. For smaller bandwidths we cannot find statistically significant coefficients.

[40] We omit the first stage of $D \times (x - c)$ in the Table 7 to save space. Full estimates can be obtained upon request to the authors.

Table 7: First Stage estimates of the Disclosure variable D for the Regression Discontinuity Design without spatial lags (bandwidths $h = 5, 6, 7, 10, 20, 30$)

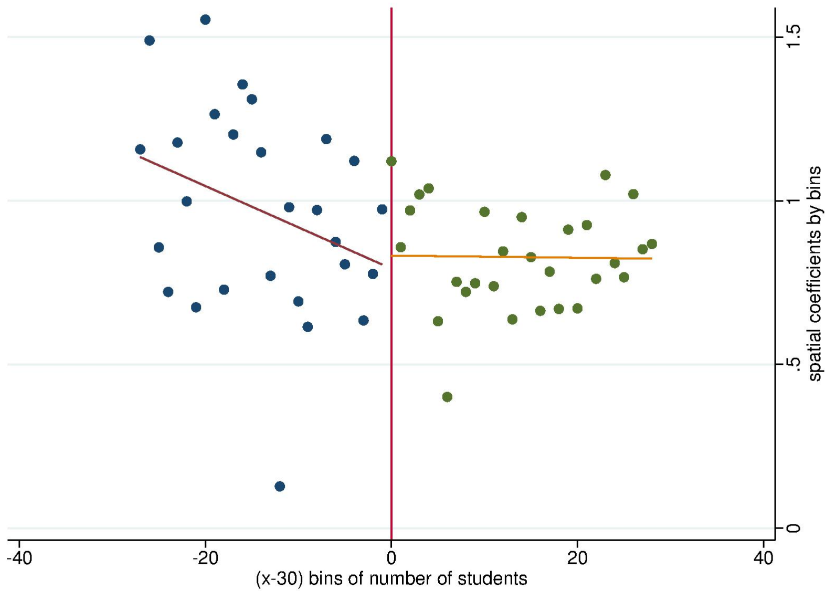
1^{st} order polynom	h=5		h=6		h=7		h=10		h=20		h=30	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
z	0.765***	0.729***	0.762***	0.681***	0.762***	0.703***	0.723***	0.697***	0.711***	0.707***	0.731***	0.727***
	(10.233)	(7.707)	(10.586)	(7.654)	(10.947)	(8.497)	(11.858)	(9.932)	(15.889)	(14.902)	(20.407)	(19.486)
$z \times (x-c)$	-0.012	-0.025	-0.010	-0.018	-0.010	-0.017	0.005	-0.005	0.010***	0.010***	0.008***	0.009***
	(-0.477)	(-0.875)	(-0.521)	(-0.819)	(-0.594)	(-0.874)	(0.480)	(-0.449)	(2.870)	(2.715)	(4.236)	(4.619)
$(x-c)$	-0.000	0.005	0.000***	0.009	0.000*	0.007	-0.000*	0.007	-0.000	-0.000	0.000	-0.001
	(-0.466)	(0.282)	(15.843)	(0.720)	(1.869)	(0.858)	(-1.824)	(1.428)	(-1.175)	(-0.109)	(1.278)	(-0.999)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	173.00	173.00	205.00	205.00	230.00	230.00	312.00	312.00	582.00	582.00	799.00	799.00
2^{nd} order polynom	h=5		h=6		h=7		h=10		h=20		h=30	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
z	0.886***	0.911***	0.848***	0.783***	0.825***	0.775***	0.800***	0.778***	0.723***	0.702***	0.698***	0.683***
	(10.275)	(6.347)	(9.994)	(6.661)	(9.927)	(7.509)	(10.364)	(8.844)	(11.603)	(10.276)	(12.990)	(11.766)
$z \times (x-c)$	-0.184**	-0.064	-0.106	-0.098	-0.071	-0.098	-0.044	-0.075*	0.006	-0.005	0.014**	0.010
	(-2.180)	(-0.481)	(-1.497)	(-1.035)	(-1.186)	(-1.291)	(-1.220)	(-1.803)	(0.472)	(-0.352)	(1.964)	(1.218)
$(x-c)$	0.034**	0.040*	0.016	0.018	0.009	0.011	0.005	0.005	0.000	-0.000	-0.000	-0.000
	(2.145)	(1.713)	(1.345)	(1.376)	(1.006)	(1.067)	(1.386)	(1.359)	(0.324)	(-0.194)	(-1.011)	(-1.319)
$z \times (x-c)^2$	-0.000	-0.081	0.000	-0.007	-0.000	0.011	-0.000	0.016	-0.000	0.008	-0.000	0.004
	(-0.029)	(-0.833)	(0.038)	(-0.125)	(-0.066)	(0.262)	(-0.246)	(0.808)	(-0.012)	(1.452)	(-0.497)	(1.029)
$(x-c)^2$	-0.000	-0.015	0.000	-0.003	-0.000	0.000	0.000	0.001	-0.000	0.000	0.000	0.000
	(-0.029)	(-0.951)	(0.027)	(-0.330)	(-0.060)	(0.085)	(0.000)	(0.532)	(-0.092)	(1.567)	(0.000)	(1.352)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	173.00	173.00	205.00	205.00	230.00	230.00	312.00	312.00	582.00	582.00	799.00	799.00

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
The model at the top panel adjusts a first degree polynomial function of $x - c$, and the model at the bottom panel adjusts a second degree polynomial funcion. h = 5; 6; 7; 10; 20; 30 represent the bandwidths around the cut-off
Theses estimates refer to one of the endogenous variable, namely the disclosure variable D. The other first stage estimates — having the interaction terms $D \times (x - c)$ and $D \times (x - c)^2$ as dependent variables— can be obtained upon request to the authors. The covariates are: gdp, wage, tax price, categorical grants, block grants, schooling, occupation rate, % men, population, % elderly, % young, % rural, % second cycle, competition, incumbent's age, incumbent's Education, left, incumbent women, aldermen's education, aldermen's age, women in council, competition for seats, fragmentation, majority of seats, president's party, governor's party, lameduck. The additional control variable $(x - c)$ represents the forcing variable — given by the number of students enrolled in the 4^{th} grade centered around the cut-off of 30 students that determines participation in Prova Brazil—, whereas $(x - c)^2$ consists of its square.
The instruments consist of the dummy variable z — that equals 1 whenever the number of enrollments in the 4^{th} grade is greater than 30 students (and 0 otherwise) — and its interaction with $(x - c)$, $z \times (x - c)$ or $z \times (x - c)^2$. Full estimates can be obtained upon request to the authors.

Table 8: Regression Discontinuity Design without spatial lags (bandwidths $h = 5, 6, 7, 10, 20, 30$)

1^{st} order polynom	h=5		h=6		h=7		h=10		h=20		h=30	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
D	0.205	-0.009	0.139	0.049	0.119	0.065	0.255*	0.120*	0.218**	0.120***	0.122	0.073*
	(1.066)	(-0.128)	(0.810)	(0.583)	(0.752)	(0.900)	(1.814)	(1.798)	(2.135)	(2.599)	(1.462)	(1.855)
$D \times (x-c)$	0.086	0.009	0.072	0.028	0.073*	0.014	0.045**	0.002	0.014**	0.005	0.002	0.002
	(1.517)	(0.398)	(1.635)	(1.453)	(1.956)	(0.813)	(2.299)	(0.226)	(1.966)	(1.534)	(0.355)	(0.830)
$(x-c)$	-0.048	0.001	-0.032	-0.010	-0.029	-0.010	-0.036***	-0.012*	-0.021***	-0.011***	-0.010***	-0.008***
	(-1.451)	(0.073)	(-1.264)	(-0.862)	(-1.391)	(-1.074)	(-2.966)	(-1.753)	(-3.694)	(-4.191)	(-2.704)	(-4.045)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	173.00	173.00	205.00	205.00	230.00	230.00	312.00	312.00	582.00	582.00	799.00	799.00
2^{nd} order polynom												
	h=5		h=6		h=7		h=10		h=20		h=30	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
D	0.397	0.173	0.453	0.160	0.381	0.134	0.117	0.055	0.207	0.081	0.327**	0.163**
	(1.388)	(1.001)	(1.540)	(0.809)	(1.379)	(0.561)	(0.603)	(0.641)	(1.356)	(1.175)	(2.250)	(2.234)
$D \times (x-c)$	1.469**	0.377	0.612*	0.074	0.311	0.130	0.165	0.041	0.085**	0.013	0.051***	0.018*
	(2.230)	(1.498)	(1.659)	(0.451)	(1.351)	(0.744)	(1.466)	(0.507)	(2.410)	(0.659)	(2.586)	(1.785)
$(x-c)$	-0.447**	-0.125	-0.280**	-0.045	-0.175*	-0.053	-0.045	-0.010	-0.045**	-0.010	-0.047***	-0.021***
	(-2.564)	(-1.504)	(-2.173)	(-0.635)	(-1.762)	(-0.570)	(-0.831)	(-0.291)	(-2.199)	(-0.871)	(-3.485)	(-2.873)
$D \times (x-c)^2$	-0.087	-0.016	0.014	0.005	0.021	-0.002	-0.010*	-0.004*	-0.001	-0.001	0.001*	0.000
	(-1.018)	(-0.644)	(0.264)	(0.492)	(0.650)	(-0.132)	(-1.833)	(-1.877)	(-0.682)	(-0.931)	(1.672)	(1.374)
$(x-c)^2$	-0.067**	-0.019	-0.036**	-0.004	-0.019	-0.005	-0.001	0.000	-0.001	0.000	-0.001***	-0.001*
	(-2.456)	(-1.592)	(-2.017)	(-0.477)	(-1.536)	(-0.456)	(-0.167)	(0.035)	(-1.254)	(0.126)	(-2.917)	(-1.888)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	173.00	173.00	205.00	205.00	230.00	230.00	312.00	312.00	582.00	582.00	799.00	799.00

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
The model at the top panel adjusts a first degree polynomial function of $x-c$, and the model at the bottom panel adjusts a second degree polynomial funcion. h = 5; 6; 7; 10; 20; 30 represent the bandwidths around the cut-off
Dependent variable y is the education spending.
Endogenous variables are the disclosure variable D and the interaction term $D \times (x-c)$. The covariates are: gdp, wage, tax price, categorical grants, block grants, schooling, occupation rate, % men, population, % elderly, % young, % rural, % second cycle, competition, incumbent's age, incumbent's Education, left, incumbent women, aldermen's education, aldermen's age, women in council, competition for seats, fragmentation, majority of seats, president's party, governor's party, lameduck. The additional control variable $(x-c)$ represents the forcing variable that is the number of students enrolled in the 4^{th} grade centered around the cut-off of 30 students that determines participation in Prova Brazil.
The instruments consist of the dummy variable z — that equals 1 whenever the number of enrollments in the 4^{th} grade is greater than 30 students (and 0 otherwise) — and its interaction with $(x-c)$, $z \times (x-c)$ and $z \times (x-c)^2$. Full estimates can be obtained upon request to the authors.



Source: Elaborated by the authors.
Figure 3: Spatial coefficients and the number of students enrolled

Figure 3 represent simple spatial correlation slopes calculated within bins of our forcing variable. We cannot observe any discontinuity in the spatial parameter through this simple visual analysis, but the number of observations by bin is very small, which makes bin analysis very noisy. However, there is a clear negative association between the forcing variable and the spatial lag parameter.

Table 9 shows the estimates of equation 7, a Regression Discontinuity Design including a spatial lag term and its interaction with the disclosure variable D. Because of the spatial lag the coefficients of the control variables cannot be interpreted as marginal effects, so they are not comparable to the estimates in Table 7. The direction of the coefficients, on the other hand, reflect that of marginal effects. We evaluate the models within the same bandwidths as before. If one gets close enough to the cutoff point one can exogenously identify the interaction between the spatial lag Wy and the disclosure variable D — DWy — on educational spending. [41] As we expect the spatial lag term to increase the explanatory power of the model the variance of the residual must be smaller. With exogenous explanatory variables around the cut-off we also expect smaller variances of the coefficients of those regressors [42]

[41] These bandwidths cannot be smaller than 5 students because otherwise, we would have to omit some control variables and instruments to gain degrees of freedom, which harms our strategy of estimation of the spatial parameter.

[42] Provided there is not strong correlation between the spatial lag and the regressors already in the model.

Models $1, 3, 5, 7, 9$ and 11 of Table 9 are reported for comparison purposes. They present the estimates of the spatial lag coefficient when we evaluate the baseline spatial model, i.e. with the effect of IDEB disclosure set to zero. Models $2, 4, 6, 8, 10$ and 12 are the ones that interest us the most in this table. For a bandwidth of 30, in model (12), the spatial correlation among the municipalities where the IDEB was not published is equal to 0.267, whereas the interaction term suggests a decrease of 0.17 correlation point among those municipalities whose IDEB was unveiled in 2008. According to Table 6 the differences between characteristics of those to the right and to the left of the cutoff is reduced in a bandwidth $h = 30$, but it is still expressive for some variables.

Restricting the bandwidth to $h = 20$, such as in model (10), eliminates part of the differences in observable and non-observable characteristics . In this case, the spatial correlation coefficient is equal to 0.255 for municipalities where IDEB was not unveiled and 0.166 correlation point smaller among those localities where the index was made public. Finally, within bandwidths $h = 5, 6$ and 7 — where the average characteristics are practically the same at both sides of the cutoff according to Table 6 — the interaction between the spatial lag and the disclosure variable presents negative and significant coefficients. Within a bandwidth of 5, model (2) shows that the spatial coefficient is estimated as 0.214 among municipalities without IDEB, whereas it is 0.116 correlation point smaller where it became public information. We can see at the bottom of the table that for the smallest bandwidth (model 2) there are only 60 degrees of freedom. If we restrict our sample to $h = 4$, the degrees of freedom drop to 30, which greatly increases the variance of the estimator.Such a result is evidence in favor of the hypothesis that information about educational quality published at the national level discourages local rulers to mimic their neighbors in terms of the level of educational spending.

At the bottom of Table 9, Hansen's statistics of overidentifying restriction and their p-values indicate that the instruments are not correlated with the error term at a 5% level of significance. Thus, the instruments are correlated with the dependent variable only through the spatial lag and should not enter the model as additional controls.

Table 9: Local Regressions with heterogeneous Spatial Lag according to the IDEB's Disclosure Status imposing first order polynomial function of $x - c$

	h=5		h=6		h=7		h=10		h=20		h=30	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Wy	0.216***	0.214***	0.181***	0.243***	0.180***	0.245***	0.162**	0.199***	0.167***	0.255***	0.160***	0.267***
	(3.191)	(3.700)	(2.757)	(4.086)	(2.682)	(3.994)	(2.345)	(2.830)	(2.764)	(4.000)	(3.064)	(4.692)
DWy		-0.116**		-0.131**		-0.131**		-0.064		-0.166***		-0.170***
		(-2.409)		(-2.376)		(-2.341)		(-1.084)		(-3.309)		(-3.852)
D		0.944**		1.080**		1.105**		0.612		1.412***		1.399***
		(2.427)		(2.377)		(2.404)		(1.269)		(3.459)		(3.932)
$D \times (x - c)$		0.024**		0.019**		0.005		-0.001		0.002		-0.002
		(2.533)		(2.197)		(0.640)		(-0.214)		(0.643)		(-1.512)
$x - c$	0.005	-0.003	0.005*	-0.003	-0.000	-0.004	-0.003*	-0.007**	-0.003***	-0.006***	-0.003***	-0.002*
	(1.470)	(-0.572)	(1.835)	(-0.721)	(-0.111)	(-1.107)	(-1.861)	(-2.240)	(-4.494)	(-3.320)	(-6.840)	(-1.696)
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
N	173.00	173.00	205.00	205.00	230.00	230.00	312.00	312.00	582.00	582.00	799.00	799.00
Degrees of Freedom	89.00	60.00	121.00	92.00	146.00	117.00	228.00	199.00	498.00	469.00	715.00	686.00
Hansen J (Overidentification)	35.77	56.54	35.98	68.14	35.77	63.75	36.25	56.50	29.41	57.59	26.85	64.80
$p > \chi^2$	0.10	0.31	0.09	0.07	0.10	0.13	0.09	0.31	0.29	0.28	0.42	0.11

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
We adjust a first degree polynomial function of $x - c$. h = 5; 6; 7; 10; 20; 30 represent the bandwidths around the cut-off
Dependent variable y is the education spending. The spatially lagged dependent variable Wy uses the contiguity criteria to assign neighborhood.
Endogenous variables are the spatially lagged dependent variable Wy, the disclosure variable D, and their interaction DWy. The covariates are: gdp, wage, tax price, categorical grants, block grants, schooling, occupation rate, % men, population, % elderly, % young, % rural, % second cycle, competition, incumbent's age, incumbent's Education, left, incumbent women, aldermen's education, aldermen's age, women in council, competition for seats, fragmentation, majority of seats, president's party, governor's party, lameduck. The additional control variable $x - c$ represents the forcing variable that is the number of students enrolled in the 4^{th} grade centered around the cut-off of 30 students that determines participation in Prova Brazil.
The instruments are the first order spatial lags of the regressors WX, the dummy variable z, that equals 1 whenever the number of enrollments in the 4^{th} grade is greater than 30 students (and 0 otherwise), interactions $z \times WX$ and $z \times (x - c)$. Full estimates can be obtained upon request to the authors.

Table 10: Local Regressions with heterogeneous Spatial Lag according to the IDEB's Disclosure Status imposing second order polynomial function of $x - c$

	h=5		h=6		h=7		h=10		h=20		h=30	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Wy	0.215***	0.184***	0.178***	0.244***	0.175***	0.231***	0.172**	0.209***	0.164***	0.250***	0.161***	0.271***
	(3.190)	(3.065)	(2.725)	(4.184)	(2.608)	(3.893)	(2.483)	(3.015)	(2.731)	(4.033)	(3.106)	(4.746)
DWy		-0.094*		-0.152***		-0.122**		-0.066		-0.155***		-0.180***
		(-1.785)		(-2.730)		(-2.252)		(-1.157)		(-3.115)		(-4.055)
D		0.784*		1.267***		0.986**		0.568		1.297***		1.476***
		(1.873)		(2.767)		(2.213)		(1.245)		(3.213)		(4.095)
$x - c$	0.005	-0.008	0.005*	-0.004	-0.000	-0.005	-0.002	-0.006**	-0.004***	-0.003	-0.003***	-0.003
	(1.445)	(-1.333)	(1.787)	(-0.895)	(-0.017)	(-1.319)	(-1.610)	(-2.039)	(-4.781)	(-1.407)	(-5.490)	(-1.633)
$D \times (x - c)$		-0.003		0.025		0.056**		0.034*		0.003		0.002
		(-0.084)		(0.860)		(2.313)		(1.827)		(0.445)		(0.369)
$(x - c)^2$	0.000	-0.002	-0.000	0.000	-0.000	-0.000	-0.000	-0.000	0.000	0.000	-0.000	-0.000
	(0.319)	(-1.613)	(-0.360)	(0.325)	(-0.659)	(-0.281)	(-1.536)	(-0.012)	(1.537)	(1.498)	(-0.257)	(-0.563)
$D \times (x - c)^2$		0.008		-0.002		-0.007**		-0.004**		-0.000		-0.000
		(1.265)		(-0.335)		(-2.070)		(-2.278)		(-1.407)		(-0.255)
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
N	173.00	173.00	205.00	205.00	230.00	230.00	312.00	312.00	582.00	582.00	799.00	799.00
Degrees of Freedom	116.00	87.00	148.00	119.00	173.00	144.00	255.00	226.00	525.00	496.00	742.00	713.00
Hansen J (Overidentification)	36.00	60.78	36.49	69.88	35.62	64.52	34.86	59.66	28.97	57.61	27.00	67.08
$p > \chi^2$	0.09	0.19	0.08	0.05	0.10	0.11	0.11	0.22	0.31	0.28	0.41	0.08

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
We adjust a second degree polynomial function of $x - c$. h = 5; 6; 7; 10; 20; 30 represent the bandwidths around the cut-off
Dependent variable y is the education spending. The spatially lagged dependent variable Wy uses the contiguity criteria to assign neighborhood.
Endogenous variables are the spatially lagged dependent variable Wy, the disclosure variable D, and their interaction DWy. The covariates are: gdp, wage, tax price, categorical grants, block grants, schooling, occupation rate, % men, population, % elderly, % young, % rural, % second cycle, competition, incumbent's age, incumbent's Education, left, incumbent women, aldermen's education, aldermen's age, women in council, competition for seats, fragmentation, majority of seats, president's party, governor's party, lameduck. The additional control variables $x - c$ and $(x - c)^2$ represents the forcing variable that is the number of students enrolled in the 4^{th} grade centered around the cut-off of 30 students that determines participation in Prova Brazil and its square.
The instruments are the first order spatial lags of the regressors WX, the dummy variable z, that equals 1 whenever the number of enrollments in the 4^{th} grade is greater than 30 students (and 0 otherwise), their interaction zWX as well as the interactions $z(x - c)$ and $z(x - c)^2$. Full estimates can be obtained upon request to the authors.

Table 10 extends the models in table 9 including a second order polynom interacting with the disclosure variable ($D \times (x - c)^2$) and as control variable ($(x - c)^2$). Introducing this additional term does not change the overall conclusion that spatial correlation is greater among the group of municipalities not subject to the publication of their IDEB. Even for the smallest bandwidth of $h = 5$ in model (2) we observe a spatial parameter 0.094 correlation point smaller among municipalities in which the IDEB was disclosed. Publishing national information about the IDEB of each municipality seems to have change the behavior of incumbents when setting the educational spending. Finally, Hansens' statistics and their p-values also suggest that instruments are not correlated with the error term at 5% significance level, and should not enter the model as control variables.

The main question to be answered is whether these findings are causal or not. With that purpose in mind a series of robustness and falsification procedures are performed in the next subsection.

6.1 Robustness and Falsification of the Results

One common test performed after RDD models consists of running it on the covariates. It is possible that a jump in the amount of grants around the cutoff of 30 students could produce either a jump in educational expenditures or a change in the spatial correlation. Table 11 present the estimates of the RDD on covariates within various bandwidths. Overall, there are no evidence of jumps on categorical or block grants. This is expected since there are not any known rule of grant distribution that distinguish municipalities with more than 30 students form those with less.

Jumps on other covariates are less plausible, but they are anyway evaluated in Table 11. Out of the control variables used in our specifications the only variable that present statistically significant jumps near the cutoff for all bandwidths is the *% women in council*. There are also some isolated cases in which we observe statistical significant estimates, but only for few bandwidths or at levels higher than 5%. There is no reason to believe this could be anything other than noise.

Table 11: Local Regression on Covariates (h = 5; 6; 7; 10; 20; 30)

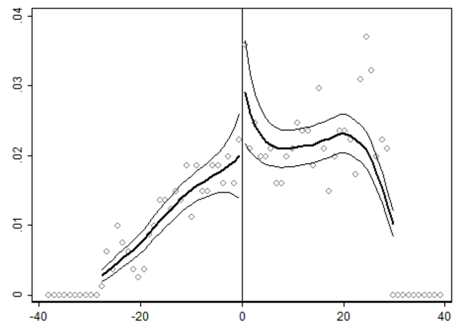
	h=5	h=6	h=7	h=10	h=20	h=30
gdp	0.237	0.229	0.104	0.210	0.159	0.049
	(0.819)	(0.871)	(0.435)	(0.992)	(1.034)	(0.449)
wage	0.087	0.068	0.062	0.088	0.015	0.012
	(0.848)	(0.737)	(0.727)	(1.143)	(0.260)	(0.283)
tax price	0.191	0.206	0.136	0.156	0.193	0.047
	(0.510)	(0.608)	(0.434)	(0.547)	(0.899)	(0.307)
categorical grants	0.000	-0.009	-0.023	0.045	0.072	0.023
	(0.003)	(-0.092)	(-0.274)	(0.607)	(1.364)	(0.608)
block grants	0.280	0.224	0.153	0.159	0.026	0.001
	(1.239)	(1.101)	(0.798)	(0.923)	(0.192)	(0.010)
schooling	0.158**	0.127*	0.111*	0.109*	0.038	0.014
	(1.971)	(1.750)	(1.656)	(1.875)	(0.961)	(0.566)
occupation	0.185	0.201	0.048	0.167	0.226	0.127
	(0.612)	(0.739)	(0.182)	(0.795)	(1.468)	(1.215)
men	0.034	0.324	0.208	-0.630	-0.809**	-0.647**
	(0.054)	(0.573)	(0.382)	(-1.215)	(-1.979)	(-2.180)
population	-0.309	-0.305	-0.219	-0.235	-0.024	0.015
	(-1.053)	(-1.176)	(-0.904)	(-1.099)	(-0.145)	(0.123)
elderly	1.602	1.095	0.587	1.374	1.331**	0.568
	(1.456)	(1.088)	(0.646)	(1.635)	(2.047)	(1.234)
young	-3.639	-2.577	-0.725	-2.118	-1.397	0.065
	(-1.369)	(-1.087)	(-0.334)	(-1.092)	(-0.930)	(0.062)
rural	-17.231	-13.223	-11.725	-12.231	-4.429	6.776
	(-1.471)	(-1.281)	(-1.206)	(-1.430)	(-0.671)	(1.365)
second cycle	2.525	5.313	-0.823	0.491	-0.289	-3.159
	(0.336)	(0.778)	(-0.127)	(0.085)	(-0.065)	(-0.951)
competition	-0.097	-0.002	0.027	0.052	0.091	0.075
	(-0.757)	(-0.019)	(0.259)	(0.563)	(1.288)	(1.449)
incumbent's age	-0.045	-0.086	-0.096	-0.090	-0.073	-0.090***
	(-0.457)	(-1.004)	(-1.258)	(-1.347)	(-1.531)	(-2.664)
incumbent's Education	0.054	0.078	0.093	0.201	-0.020	-0.028
	(0.262)	(0.424)	(0.551)	(1.326)	(-0.179)	(-0.349)
left	-0.048	-0.058	0.022	-0.015	0.045	0.063
	(-0.256)	(-0.360)	(0.144)	(-0.123)	(0.499)	(0.985)
incumbent women	-0.030	0.026	0.065	0.049	0.069	0.018
	(-0.294)	(0.274)	(0.757)	(0.692)	(1.327)	(0.440)
aldermen's education	0.793	-1.451	-1.842	-0.274	0.688	1.703
	(0.169)	(-0.354)	(-0.501)	(-0.082)	(0.264)	(0.892)
aldermen's age	0.053	0.042	0.021	0.025	0.033	0.015
	(1.539)	(1.308)	(0.709)	(0.951)	(1.608)	(1.006)
women in council	-10.202**	-9.946**	-8.863**	-10.447***	-4.669*	-3.036*
	(-2.065)	(-2.308)	(-2.248)	(-2.920)	(-1.810)	(-1.672)
competition for seats	-0.144	-0.065	0.019	-0.037	0.098	0.180**
	(-0.742)	(-0.385)	(0.117)	(-0.270)	(0.932)	(2.359)
fragmentation	3.187	3.236	5.680	1.188	1.066	2.483
	(0.635)	(0.750)	(1.427)	(0.342)	(0.434)	(1.420)
majority of seats	0.327	0.312*	0.149	0.150	0.082	0.021
	(1.599)	(1.709)	(0.876)	(1.009)	(0.748)	(0.264)
president's party	-0.011	0.004	0.053	0.052	0.063	0.052
	(-0.090)	(0.038)	(0.586)	(0.660)	(1.128)	(1.294)
governor's party	-0.023	0.094	0.015	0.042	-0.001	0.076
	(-0.204)	(0.800)	(0.143)	(0.420)	(-0.014)	(1.235)
lameduck	0.003	0.007	-0.072	-0.052	-0.112	-0.094
	(0.016)	(0.042)	(-0.470)	(-0.389)	(-1.109)	(-1.297)

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
The second stage of the Parametric Regression Discontinuity Designs ran on covariates consists of $X_k = \alpha + \beta D + \gamma D \times (x-c) + \theta(x-c) + \epsilon$ where X_k is kth, D is the IDEB publishing indicator, $(x-c)$ is the forcing variable.have the following specification. The first stage consists of runing D and $D \times (x-c)$ on the instruments z and $z \times (x-c)$ and the number of students centerer around the cutoff $(x-c)$.

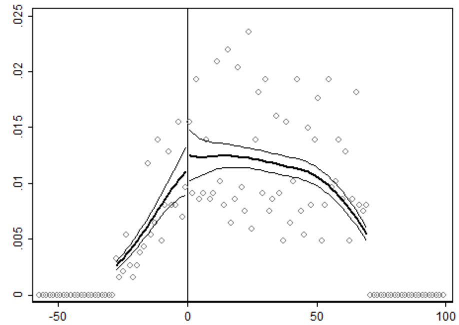
We run McCrary (2008) tests to check for evidence of manipulation in the running variable. We perform the test for different intervals around the cutoff. The McCrary density test requires the definition of binwidths and bandwidths. Since there is not consensus about which are the best sizes, we let the data-based algorithm described in McCrary (op cit.) choose them for us. We only vary the total amount of data used around the cutoff, which is very asymmetrically distributed around that point. Starting with a window of 30 students to the right and 30 students to the left of the cutoff point, we increase the amount of data by increasing the interval to the right to 70 students, and then to 170 and 270 students, respectively. Using these subsets of data prevents the algorithm of choosing bandwidths that are too big, which does not make sense in case one have few observations at the left side of the cutoff —as in our case. The graphs are shown in figure 4. A quick inspection of the graphs shows no evidence of manipulation of the running variable around the cutoff. In the first graph a small jump in the density is most likely the result of the outlier next to the cutoff. But the confidence interval on each side coincide, indicating no statistical significance in the jump. When we increase the amount of observations, the optimal binwidths and bandwidths get bigger, and in neither case we observe jumps in the forcing variable.

Finally Tables 12 and 13 show falsification tests for both linear and quadratic models in Tables 9 and 10. The falsification consists of testing the effect of the disclosure variable on the education expenditure before the actual disclosure. We should not find statistically significant coefficients, otherwise, there is evidence of spurious association between the regressors and the regressand. We evaluate the spatial parameters and their interaction using data from 2003 to 2007 (thus, prior to 2008), and data from 2007 alone. As the publication of IDEB happened in 2007, we want to show that there is no immediate effects because the budget of 2007 could not be amended in the same year of the IDEB publication.

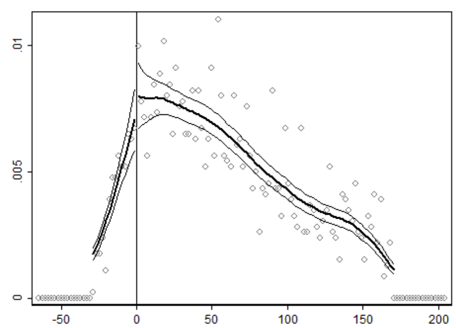
In Table 12, within the greater bandwidth of 30 students and data from 2007 we find a spatial correlation of 0.225 among municipalities where IDEB was not published later in 2008, and a coefficient 0.215 correlation point smaller where the index was published . For the period between 2003 and 2007, we find a spatial coefficient equal to 0.303 between municipalities without IDEB, which is reduced by 0.215 correlation point in municipalities that had their IDEB disclosed in 2008. For a bandwidth of 20 students and data between 2003 and 2007 — or 2007 alone— we still find smaller spatial coefficients where IDEB was published (municipalities to the right of the cutoff). This means the results are not causal for those bandwidths. The naive analysis in Figure 3 had anticipated this finding by showing a negative association between the spatial coefficients and the forcing variable. Thus, within



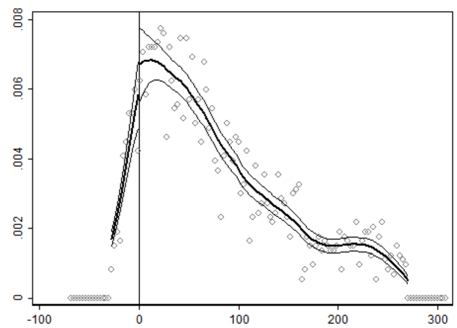
(a) $w \in [x - c - 30; \quad x - c + 30]$
optimal bin=1.04, optimal bandwidth=10.09



(b) $w \in [x - c - 30; \quad x - c + 70]$
optimal bin=1.35, optimal bandwidth=28.39



(c) $w \in [x - c - 30; \quad x - c + 170]$
optimal bin=2.12, optimal bandwidth=35.12



(d) $w \in [x - c - 30; \quad x - c + 270]$
optimal bin=2.89, optimal bandwidth=38.3

Figure 4: McCrary (2008) density test

big bandwidths we would get smaller spatial coefficients to the right of any cutoff, which makes the reduction of the bandwidth mandatory.

On the other hand, for the smaller bandwidths of 5, 6, 7 and 10 students, we do not find any statistically significant difference in the spatial coefficients of the groups with or without IDEB. The coefficients on the disclosure variable D alone are not significant either. In the year of 2007, within the smaller bandwidth of $h = 5$, the estimate of the spatial coefficient is equal to 0.09 and the coefficient on spatial interaction DWy is not significant. We obtain the same results for bandwidths $h = 6, 7$ and 10. With the data covering the entire period between 2003 and 2007 the spatial coefficient for $h = 5$ is 0.23, whereas the coefficient on the interaction between the spatial lag and the disclosure variable DWy is also non significant. Bandwidths $h = 6, 7$ and 10 show the same pattern, i.e. no statistically significant difference between the spatial correlation of municipalities with and without IDEB. Table 13 confirm such results for a quadractic function of the forcing variable $(x - c)$. Thus, around the cutoff, i.e. for $h - 5, 6, 7$ and 10 our estimates in Tables 9 and 10 can be claimed as causal. This reinforces our hypothesis that more public information reduces (both information asymmetry and) yardstick competition.

Table 12: Falsification test for previous years imposing first degree polynomial function of $x-c$ and bandwidths h = 5; 6; 7; 10; 20; 30 around the cut-off

	h=5		h=6		h=7		h=10		h=20		h=30	
	(2)		(4)		(6)		(8)		(10)		(12)	
	2007	2003-07	2007	2003-07	2007	2003-07	2007	2003-07	2007	2003-07	2007	2003-07
Wy	0.097*	0.230***	0.104*	0.197***	0.090	0.202***	0.121*	0.199***	0.193***	0.265***	0.225***	0.303***
	(1.811)	(5.833)	(1.843)	(5.022)	(1.434)	(5.202)	(1.916)	(5.415)	(3.199)	(8.830)	(4.347)	(10.958)
DWy	0.044	0.035	0.018	0.009	0.005	-0.009	-0.031	-0.037	-0.214***	-0.201***	-0.285***	-0.215***
	(1.081)	(1.185)	(0.379)	(0.286)	(0.097)	(-0.298)	(-0.594)	(-1.310)	(-4.583)	(-8.825)	(-6.675)	(-10.295)
D	-0.344	-0.261	-0.133	-0.081	-0.017	0.065	0.237	0.268	1.717***	1.607***	2.252***	1.679***
	(-1.026)	(-1.118)	(-0.339)	(-0.329)	(-0.039)	(0.266)	(0.558)	(1.190)	(4.669)	(8.910)	(6.695)	(10.270)
$D \times (x-c)$	0.021*	0.014*	0.028**	0.018**	0.010	0.009	-0.006	0.002	0.004	0.007***	0.001	0.003**
	(1.705)	(1.693)	(2.573)	(2.508)	(0.959)	(1.533)	(-0.969)	(0.593)	(1.379)	(4.300)	(0.419)	(2.569)
$x-c$	-0.011**	-0.014***	-0.011**	-0.011***	-0.010***	-0.010***	-0.003	-0.005**	-0.008***	-0.010***	-0.005***	-0.006***
	(-2.204)	(-3.168)	(-2.287)	(-2.720)	(-2.617)	(-2.976)	(-0.929)	(-2.389)	(-3.888)	(-8.319)	(-2.961)	(-6.100)
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
N	173.00	865.00	205.00	1025.00	230.00	1150.00	312.00	1560.00	582.00	2910.00	799.00	3995.00

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
We adjust a first degree polynomial function of $x-c$. h = 5; 6; 7; 10; 20; 30 represent the bandwidths around the cut-off
Dependent variable y is the education spending. The spatially lagged dependent variable Wy uses the contiguity criteria to assign neighborhood.
Endogenous variables are the spatially lagged dependent variable Wy, the disclosure variable D, and their interaction DWy. The covariates are: gdp, wage, tax price, categorical grants, block grants, schooling, occupation rate, % men, population, % elderly, % young, % rural, % second cycle, competition, incumbent's age, incumbent's Education, left, incumbent women, aldermen's education, aldermen's age, women in council, competition for seats, fragmentation, majority of seats, president's party, governor's party, lameduck. The additional control variables $x-c$ and $(x-c)^2$ represents the forcing variable that is the number of students enrolled in the 4^{th} grade centered around the cut-off of 30 students that determines participation in Prova Brazil and its square.
The instruments are the first order spatial lags of the regressors WX, the dummy variable z, that equals 1 whenever the number of enrollments in the 4^{th} grade is greater than 30 students (and 0 otherwise), their interaction zWX as well as the interaction $z(x-c)$. Full estimates can be obtained upon request to the authors.

Table 13: Falsification test for previous years imposing second degree polynomial function of $x-c$ and bandwidths h = 5; 6; 7; 10; 20; 30 around the cut-off

	h=5		h=6		h=7		h=10		h=20		h=30	
	(2)		(4)		(6)		(8)		(10)		(12)	
	2007	2003-07	2007	2003-07	2007	2003-07	2007	2003-07	2007	2003-07	2007	2003-07
Wy	0.090*	0.198***	0.130**	0.186***	0.103	0.207***	0.142**	0.197***	0.212***	0.272***	0.224***	0.316***
	(1.704)	(4.966)	(2.242)	(4.696)	(1.610)	(5.289)	(2.271)	(5.343)	(3.634)	(9.175)	(4.237)	(11.059)
DWy	0.042	0.055*	0.003	0.019	0.005	-0.007	-0.029	-0.032	-0.202***	-0.190***	-0.287***	-0.228***
	(1.084)	(1.785)	(0.060)	(0.590)	(0.093)	(-0.209)	(-0.556)	(-1.137)	(-4.347)	(-8.356)	(-6.738)	(-10.609)
D	-0.322	-0.362	0.037	-0.064	-0.014	0.083	0.236	0.246	1.651***	1.509***	2.278***	1.799***
	(-1.014)	(-1.513)	(0.095)	(-0.253)	(-0.032)	(0.333)	(0.556)	(1.093)	(4.476)	(8.327)	(6.769)	(10.565)
$x-c$	-0.019***	-0.025***	-0.014***	-0.016***	-0.010**	-0.010**	-0.001	-0.004	-0.003	-0.006***	-0.005**	-0.011***
	(-3.215)	(-4.634)	(-2.675)	(-3.595)	(-2.282)	(-2.339)	(-0.179)	(-1.534)	(-1.073)	(-3.358)	(-1.998)	(-6.541)
$D\times(x-c)$	0.007	-0.033	-0.026	-0.063***	0.009	-0.020	-0.005	-0.007	-0.011	0.001	-0.001	0.009***
	(0.225)	(-1.238)	(-0.825)	(-2.743)	(0.294)	(-1.098)	(-0.250)	(-0.541)	(-1.355)	(0.172)	(-0.259)	(3.110)
$(x-c)^2$	-0.004**	-0.004***	-0.001	-0.001	-0.000	0.000	0.001	0.000	0.000***	0.000**	0.000	-0.000***
	(-2.107)	(-2.694)	(-0.959)	(-1.094)	(-0.087)	(0.516)	(1.202)	(0.516)	(2.806)	(2.214)	(0.312)	(-2.983)
$D\times(x-c)^2$	0.009	0.016***	0.010**	0.015***	0.000	0.004	-0.001	0.001	-0.000	-0.000	0.000	0.000*
	(1.507)	(3.306)	(2.105)	(4.023)	(0.065)	(1.430)	(-0.700)	(0.562)	(-0.200)	(-0.815)	(0.110)	(1.842)
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
N	173.00	865.00	205.00	1025.00	230.00	1150.00	312.00	1560.00	582.00	2910.00	799.00	3995.00

Notes: t-statistics in parenthesis; *** significant at 1%; ** significant at 5%; * significant at 10%.
We adjust a second degree polynomial function of $x-c$. h = 5; 6; 7; 10; 20; 30 represent the bandwidths around the cut-off
Dependent variable y is the education spending. The spatially lagged dependent variable Wy uses the contiguity criteria to assign neighborhood.
Endogenous variables are the spatially lagged dependent variable Wy, the disclosure variable D, and their interaction DWy. The covariates are: gdp, wage, tax price, categorical grants, block grants, schooling, occupation rate, % men, population, % elderly, % young, % rural, % second cycle, competition, incumbent's age, incumbent's Education, left, incumbent women, aldermen's education, aldermen's age, women in council, competition for seats, fragmentation, majority of seats, president's party, governor's party, lameduck. The additional control variables $x-c$ and $(x-c)^2$ represents the forcing variable that is the number of students enrolled in the 4^{th} grade centered around the cut-off of 30 students that determines participation in Prova Brazil and its square.
The instruments are the first order spatial lags of the regressors WX, the dummy variable z, that equals 1 whenever the number of enrollments in the 4^{th} grade is greater than 30 students (and 0 otherwise), their interaction zWX as well as the interactions $z(x-c)$ and $z(x-c)^2$. Full estimates can be obtained upon request to the authors.

7 Concluding Remarks

Yardstick competition arises from asymmetric information between voters and incumbents. The latter are better informed, whereas the former are imperfectly informed. To better choose their local rulers, voters assess their neighbors' policies and take them as a benchmark. Incumbents whose policies are relatively better will be reappointed and those that are relatively worse than their neighbors will not. This forces incumbents to "mimic" their neighbors to not fall behind them.

We explore the nationwide public release of the Brazilian Basic Education Development Index for Brazilian municipalities in 2007 to estimate whether spatial strategic interactions among those municipalities have decreased after the disclosure. The preliminary analysis in this paper relies on a panel of Brazilian municipalities (from 2003-2011) to test whether local-level disclosure of the IDEB in mid-2007 diminished the spatial interaction among jurisdictions in terms of educational spending, which could be attributed to a reduction in the information asymmetries regarding educational quality. The results suggest that spatial correlation decreased by 0.124 (20% of the total spatial correlation) correlation point after the IDEB was disclosed. Although important,this "before and after" analysis may be biased if another important event has occurred concomitantly with the publishing of the IDEB, confounding with the effects of the later.

In our main exercise we take advantage of IDEB's discontinuity in the enrollment-based rule that determines the unveiling of the IDEB of those municipalities with more than 30 students. As one approaches the cutoff, the differences in terms of non-observable characteristics correlated with the disclosure dummy vanish and we can identify the effects of the IDEB disclosure on the spatial interaction pattern. So, we restrict our analysis to bandwidths close enough to the cutoff to find causal evidence linking the IDEB disclosure and yardstick competition. We conduct McCrary tests to check whether the cutoff rule has been manipulated, but we do not find evidence in this sense. We also run RDD on covariates to check whether there are "jumps" in these variables around the cut-off, but we also can not find evidence in this direction. Finally, we perform falsification tests evaluating the effects of the disclosure of the IDEB on the past educational expenditure — regarding the previous period between 2003 and 2007, or 2007 alone— and on the spatial coefficient. The tests' results indicate that only for the smaller bandwidths of $h = 5, 6, 7$ and 10 we can make causal statements about the effects of IDEB disclosure status on the education spending or on the spatial parameter.

Our estimate for a bandwidth of 5 students around the cutoff suggest that the IDEB disclosure reduces spatial correlation in 0.116 correlation point (54% of the observed spatial correlation). Similar results are found for bandwidths $h = 6$ and 7, with reductions of 0.131 correlation point for both bandwidths. Inserting a quadratic polynom in $(x - c)$ does not change our conclusion. For $h = 5$ spatial correlation drops 0.094 correlation points (51% of the observed spatial correlation) among municipalities where IDEB were published. The same conclusions can be drawn for estimates within bandwidths $h = 6$ and 7, where spatial correlation are respectively 0.152 and 0.122 correlation point smaller among municipalities without IDEB. Overall, the results reinforce our hypothesis that more information on educational quality reduces information asymmetry and consequently, reduces yardstick competition.

As Revelli (2006) notes, if there is institutional change that reduces information asymmetry, the yardstick competition will be discouraged, thus reducing spatial interaction. This paper aims to complement this literature by establishing the relationship between performance evaluation, accountability and yardstick competition in the specific case of educational spending. In addition, considering the uniqueness of the Brazilian framework, the results reinforce the external validity of the phenomenon.

References

Alesina, A. & Cukierman, A. (1990), 'The Politics of Ambiguity', *Quarterly Journal of Economics* **105**(4), 829–850.

Alesina, A. & Spear, S. (1988), 'An overlapping generations model of electoral competition', *Journal of Public Economics* **37**(3), 359–379.

Allers, M. & Elhorst, P. (2005), 'Tax Mimicking and Yardstick Competition Among Local Governments in the Netherlands', *International Tax and Public Finance* **12**(4), 493–513.

Alt, J. & Lowry, R. (1994), 'Divided Government, Fiscal Institutions, and Budget Deficits: Evidence from the States', *The American Political Science Review* **88**(4), 811–828.

Anselin, L., Bera, A., Florax, R. & Yoon, M. (1996), 'Simple diagnostic tests for spatial dependence', *Regional Science and Urban Economics* **26**(1), 77–104.

Arvate, P. & Zoghbi, A. (2010), 'Intergenerational conflict and public education expenditure when there is co-residence between the elderly and young', *Economics of Education Review* **29**(6), 1165–1175.

Bergstrom, T. C., Rubinfeld, D. L. & Shapiro, P. (1982), 'Micro-Based Estimates of Demand Functions for Local School Expenditures', *Econometrica* **50**(5), 1183–1205.

Besley, T. & Case, A. (1995), 'Incumbent behavior: vote seeking, tax-setting and yardstick competition', *The American Economic Review* **85**(1), 25–45.

Besley, T. & Smart, M. (2007), 'Fiscal restraints and voter welfare', *Journal of Public Economics* **91**(3-4), 755–773.

Bivand, R. & Szymanski, S. (2000), 'Modelling the spatial impact of the introduction of compulsive competitive tendering', *Regional Science and Urban Economics* **30**(2), 203–219.

Borcherding, T. E. & Deacon, R. T. (1972), 'The Demand for the Services of Non-Federal Governments', *The American Economic Review* **62**(5), 891–901.

Bordignon, M., Cerniglia, F. & Revelli, F. (2003), 'In search of yardstick competition: a spatial analysis of Italian municipality property tax-setting', *Journal of Urban Economics* **54**(2), 199–217.

Bordignon, M., Cerniglia, F. & Revelli, F. (2004), 'Yardstick competition in intergovernmental relationships: theory and empirical predictions', *Economics Letters* **83**(3), 325–333.

Brueckner, J. (2003), 'Strategic interaction among governments: an overview of empirical studies', *International Regional Science Review* **26**(2), 175–188.

Card, D. & Payne, A. (2002), 'School finance reform, the distribution of school spending, and the distribution of student test scores', *Journal of Public Economics* **83**(1), 49–82.

Case, A., Hines, J. & Rosen, H. (1993), 'Budget spillovers and fiscal policy interdependence: Evidence from the States', *Journal of Public Economics* **52**(3), 285–307.

Desposato, S. (2006), 'Parties for Rent? Ambition, Ideology, and Party Switching in Brazil's Chamber of Deputies', *American Journal of Political Science* **50**(1), 62–80.

Elhorst, P. (2010), 'Applied Spatial Econometrics: Raising the Bar', *Spatial Economic Analysis* **5**(1), 9–28.

Elhorst, P. & Fréret, S. (2009), 'Evidence of political yardstick competition in France using a two-regime spatial Durbin model with fixed effects', *Journal of Regional Science* **49**(5), 931–951.

Ferejohn, J. (1986), 'Incumbent Performance and Electoral Control', *Public Choice* **50**(1-3), 5–25.

Ferrazt, C. & Finan, F. (2008), 'Evidence of political yardstick competition in France using a two-regime spatial Durbin model with fixed effects', *The Quarterly Journal of Economics* **123**(2), 703–745.

Firpo, S., Pieri, R. & Souza, A. (2011), Electoral Impacts of Quality Improvements in Basic Education: Evidence from Brazilian Municipalities, *in* 'Annals of the 33rd meeting of the Brazilian Econometric Society', pp. 1–31.

Geys, B. (2006), 'Looking across borders: a test of spatial policy interdependence using local government efficiency ratings', *Journal of Urban Economics* **60**(3), 443–462.

Hanushek, E. (1986), 'The Economics of Schooling: Production and Efficiency in Public Schools Reviewed', *Journal of Economic Literature* **24**(3), 1141–1177.

Hanushek, E. (1996), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, first edn, Brookings Institution Press, Washington, chapter 2, pp. 43–73.

Hanushek, E. (2006), *Handbook of the Economics of Education*, Vol. 2, first edn, Elsevier, Washington, chapter 14, pp. 865–908.

Ladd, H. (1992), 'Mimicking of local tax burdens among neighboring counties', *Public Finance Quarterly* **20**(4), 450–467.

Lee, L. (2004), 'Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models', *Econometrica* **72**(6), 1899–1925.

Leuven, E., Lindahl, M., Oosterbeek, H. & Webbink, D. (2007), 'The effect of extra funding for disadvantaged pupils on achievement', *Review of Economics and Statistics* **89**(4), 721–736.

Lucas, K. & Samuels, D. (2010), 'The Ideological Coherence of the Brazilian Party System, 1990-2009', *Journal of Politics in Latin America* **2**(3), 39–69.

Menezes-Filho, N. A. & Pazzello, E. (2007), 'Does money in schools matter? Evaluating the effects of a funding reform on wages and test scores in Brazil', *Economics of Education Review* **26**(6), 660–672.

Milyo, J. & Schosberg, S. (2000), 'Gender bias and selection bias in House elections', *Public Choice* **105**(1-2), 41–59.

Mukherjee, B. (2003), 'Political parties and the size of government in multi party legislatures: examining cross-country and panel data evidence', *Comparative Political Studies* **36**(6), 699–728.

OECD (2009), Equally prepared for life?: How 15-year-old boys and girls perform in school, Technical report, The Organization for Economic Co-operation and Development, Brasília.
URL: *http://www.oecd.org/dataoecd/59/50/42843625.pdf*

Persson, T., Roland, G. & Tabellini, G. (1997), 'Separation of Powers and Political Accountability', *The Quarterly Journal of Economics* **112**(4), 1163–1202.

Poterba, J. M. (1997), 'Demographic structure and the political economy of public education', *Journal of Policy Analysis and Management* **16**(1), 48–66.

Revelli, F. (2001), 'Spatial patterns in local taxation: tax mimicking or error mimicking?', *Applied Economics* **33**(9), 1101– 1107.

Revelli, F. (2002), 'Local taxes, national politics and spatial interactions in English district election results', *European Journal of Political Economy* **18**(2), 281–299.

Revelli, F. (2006), 'Performance rating and yardstick competition in social service provision', *Journal of Public Economics* **90**(3), 458–474.

Revelli, F. (2008), 'Performance competition in local media markets', *Journal of Public Economics* **92**(7), 1585–1594.

Revelli, F. (2009), 'Spend more, get more? An inquiry into English local government performance', *Oxford Economic Papers* **62**(1), 185–207.

Revelli, F. & Tovmo, P. (2007), 'Revealed yardstick competition: local government efficiency patterns in Norway', *Journal of Urban Economics* **62**(1), 121–134.

Rincke, J. (2009), 'Yardstick competition and public sector innovation', *International Tax and Public Finance* **16**(3), 337–361.

Salmon, P. (1987), ‘Decentralization as an Incentive Scheme’, *Oxford Review of Economic Policy* **3**(2), 24–43.

Sollé Ollé, A. (2003), ‘Electoral Accountability and Tax Mimicking the Effects of Electoral Margins, Coalition, Government Ideology’, *European Journal of Political Economy* **19**(4), 685–713.

Sollé Ollé, A. (2006), ‘The effects of party competition on budget outcomes: Empirical evidence from local governments in Spain’, *Public Choice* **126**(1), 145–176.

Strömberg, D. (2004), ‘Mass Media Competition, Political Competition, and Public’, *The Review of Economic Studies* **71**(1), 265–284.

Weingast, B., Shepsle, K. & Johnsen, C. (1981), ‘The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics’, *The Journal of Political Economy* **89**(4), 642–664.