Political Learning and Officials' Motivations: An Empirical Analysis of the Education Reform in the State of São Paulo

Thomaz Gemignani* Ricardo Madeira*

Abstract

We investigate the occurence of social learning among government officials in a context of decentralization of political responsibilities - the schooling decentralization reform of the state of São Paulo - and use it to analyze officials' motivations driving the adhesion to that program. We explore how the information exchange about the newly adopted tasks is configured and which aspects about the returns of decentralization are mostly valued by officials in their learning process. In particular, we try to determine to what extent the adhesion to the reform was due to electoral motivations or, rather, to concerns about the quality of public education provision. We present evidence that social learning configures a relevant factor in the reform implementation and find that mayors are more likely to adhere to the program upon the receipt of good news about the electoral returns of decentralization. On the other hand, experiences by information neighbors that turn out to be successful in improving the public provision seem to be ignored in mayors' decisions for decentralization. The argument for electoral motivations is further supported by evidence that officials tend to be more responsive to information transmitted by neighbors affiliated to the same party as their own.

JEL Classification: D78, D83, I28.

Keywords: Electoral Incentives; Learning; Decentralization.

*Faculdade de Economia, Administração e Contabilidade - Universidade de São Paulo.

1 Introduction

In the recent decades, developing economies have resorted to a diverse set of policy experiments in order to improve efficiency in the public provision of services. Regrettably, the conception of such policies usually fails to take into account the incentives of the agents in charge of their conduction, resulting in the waste of public resources and possibly in adverse results unforeseen by policymakers. In this sense, unraveling officials' motivations behind the choice among alternative policies is of great importance for an appropriate formulation of public policy.

One of the most prevalent options by such economies has been the decentralization of political responsibilities, in which not only the transfer of administrative functions (and corresponding financial resources) but also the attribution of political power of decision making - subject to a mechanism of public choice - is made towards lower levels of government. Firstly, central governments are usually associated with a more inefficient operation and higher levels of corruption, while local governments would detain an informational advantage over preferences and needs of local communities due to their geographic proximity and lower costs of monitoring[1] (Oates (1999)). Furthermore, as pointed out in Seabright (1996), decentralization in government may be attractive for its promotion of a higher accountability to local preferences and the subsequent provision of correct incentives to local officials, which are induced to make effective use of any information advantages they may possess.[2] Also, these reforms may improve efficiency by introducing in democratic contexts an element of higher interjurisdictional competition.[3]

On the other hand, developing countries are tipically characterized by a lower quality of local administrative teams (Bardhan (2002)) and may be subject to inferior information regarding the necessary procedures for an adequate provision of the service under decentralization, due to their usual lack of experience with the newly imposed responsibilities. Clearly, such lower knowledge may gravely compromise the quality of the decentralized provision and the officials' electoral performance. In particular, in those cases where local governments may choose to adhere to the decentralization reform, its progress may be severed by this inferior know-how. Hence, while simultaneously presenting several agents with new technologies and/or obligations, iniciatives of decentralization of public services constitute events

[1] As argued in Oates (1999), the expectation that central levels of government present a more uniform provision among jurisdictions may be due not only to the reasons cited above, but also to the existence of political pressures and constitutional constraints that limit its capacity of providing distinct levels of goods.

[2] In fact, the relevance of the concept of accountability in decentralization analyses is reinforced by the observation that it may induce a better performance by the government even when jurisdictions present homogeneous preferences

[3] See, for example, Tiebout (1956) and Epple & Zelenitz (1981).

propitious to the emergence of social learning.[4]

The role of learning by social interactions has been widely studied in several strands of economic literature, such as in endogenous growth (e.g. Aghion & Howitt (1998)) and in urbanization (e.g. Glaeser et al. (1992)). Some evidence regarding the notion that imperfect knowledge about aspects of an introduced technology may configure a barrier to its adoption and to productivity improvements are presented, for instance, by Foster & Rosenzweig (1995) and Conley & Udry (2010) in the context of agricultural technology diffusion. Notwithstanding, we have no knowledge of studies adressing the question of how these issues may arise in political contexts and to what degree they may compromise political performance. In the present analysis, we investigate the ocurrence of social learning in the school decentralization reform implemented in the state of São Paulo and use it to explore the officials' motivations behind the adhesion to that program. More specifically, we try to determine to what extent the adhesion to the decentralization program was due to electoral motivations or, rather, to genuine concerns about the quality of education provision.

A key aspect of such policy was that mayors (local governments) were able to choose their degree of adhesion to the reform in a rather diversified manner: each mayor could decide for the adoption of each state school in his or her municipality separately (on a school-by-school basis) and at any time, which granted the reform a gradual implementation. Moreover, municipal administrations had virtually no experience regarding the provision of the levels of schooling under decentralization, and were thus prone to the exchange of information about the impacts of adhesion to the program. We define alternative measures of success in school adoption experiences (regarding separately the electoral and the provision quality dimensions) and investigate the role of information neighbors' successful experiences in the mayors' decision making over adhesion to the program, as well as the kind of information mostly valued by the local officials. That is, we wish to explore whether the finding that school adoption gives rise to a better electoral performance plays a greater role in motivating mayors towards adhesion than the finding that adoption enhances social welfare by favouring a better provision of education services. Of course, elucidating the relative importance of these alternative channels in the politicians' decision making is an essential matter not only in educational contexts, but also to the appropriate design of public policies in general.

Through the use of data on exogenously defined information networks, we find evidence that the receipt of good news in the electoral sense by information neighbors tends to increase the adhesion of mayors to the de-

[4] Alternatively, the presence of yardstick competition in decentralization scenarios may induce more frequent trials of alternative inputs in search of more efficient methods of provision, and thus promote a better learning.

centralization program (i.e. the probability that a mayor adopts a given state school in its municipality). On the other hand, mayors seem to be unafected from learning about successful experiences in the quality dimension. Moreover, our suggestion that mayors' decisions over decentralization is primarily electorally motivated is further supported by evidence that mayors tend to be even more responsive to the information transmitted by those neighbors affiliated to the same political party as their own.

This paper is organized as follows. Section 2 describes the decentralization process under examination and section 3 describes our main data sources and the construction of information networks. Section 4 presents our empirical strategy. Our main results are described in section 5 and robustness checks and extensions are considered in section 6. Finally, section 7 presents our conclusions.

2 Institutional Background

2.1 Primary Education Funding and Decentralization in Brazil

The Brazilian pre-college educational system is organized into four levels: preschool (attended by 6 year-olds), primary school (attended by 7 to 10 year-olds), secondary school (attended by 11 to 14 year-olds) and high school (attended by 15 to 17 year-olds). Except for preschool, the public provision of these educational levels was typically the responsibility of the states' administrations until the late 1990's.

While the first steps of the process of decentralization of public education provision towards municipalities were induced by the Federal Constitution of 1988, education funding regulation was not well established until almost a decade later. In order to ensure the fulfilment of the municipalities' newly imposed obligations, the federal government implemented a national education bill (known as FUNDEF), approved by the national congress in 1996 (but not implemented until January 1998), which played a major role in shaping the public resources earmarked to education.[5]

The Brazilian Constitution mandates that municipalities and states must spend at least 25% of their tax revenues and transfers on their educational systems. However, until the FUNDEF implementation, there was no regulation about how these resources should be spent or how they should be distributed across various levels of education. Due to the heterogeneity between states and municipalities with respect to the number of pupils enrolled in their educational systems, richer states and municipalities were spending more per pupil than their poorer counterparts. In addition, richer states

[5]FUNDEF stands for *Fundo para a Manutenção e Desenvolvimento do Ensino Fundamental e Valorização do Magistério.*

and municipalities could exploit the broad definition of educational expenditures and spend their resources in activities only marginally related to education. The lack of effective monitoring also contributed to this type of moral hazard behavior.

The essential feature of the FUNDEF was the creation of a fund with resources collected from states and municipalities. Each state/municipality must contribute 15% of its tax revenues and transfers and collected resources would be redistributed according to the number of students enrolled in their primary and secondary education systems. Also, a commission was appointed to annually set a minimum monetary value per pupil (based on an estimate of school management costs) to be sustained by the federal government when necessary. Finally, states and municipalities should spend at least 60% of the received resources on teacher wages.

Thus, the funding reform embodied in the FUNDEF exerted a relevant influence over the incentives underlying the decentralization of schooling: in order to retrieve some of the resources collected by the fund to themselves, municipalities should then proceed to further develop (or institute) their primary and secondary education systems. As argued below, this mechanism was particularly essential to the decentralization process established in the state of São Paulo.

2.2 The Decentralization Program of the State of São Paulo

In 1995, the Brazilian Social Democratic Party (PSDB) took the government office of the state of São Paulo, after winning the 1994 election. There was consensus among elected officials on the inefficiency of the state's highly centralized educational system[6], stemming from two key observations. First, the existence of numerous bureaucratic tiers between state government policy makers and schools' principals, which imposed impediments to the system's response to the schools' specific needs. Second, the lack of community involvement in local school management. In order to tackle these problems, the state government launched one of the largest decentralization programs ever implemented in the Brazilian public education system, known as *Municipalização do Ensino*. In particular, the reform was expected to make educational policy decision-making more responsive to the local communities' preferences by making use of the higher accountability of municipal governments (relative to the state government) to the desires of any particular locality.[7] Moreover, decentralization could increase the participation of local communities simply by bringing policy makers geographically closer to the citizens, thus further improving the response of school management to the communities' needs. These arguments were particularly appealing to

[6]In 1997, São Paulo had the third highest proportion of students enrolled in state schools (vis-à-vis municipal schools) among all the 27 Brazilian states.

[7]See, for example, Seabright (1996).

the São Paulo state due to the huge social and economic differences across its various regions.

Differently from most decentralization programs examined in the literature, in the São Paulo state reform the decision to decentralize was also decentralized. That is, the state government devolved to each municipality the decision to take over the primary and secondary state schools located within their jurisdiction. The municipalities were allowed to make this decision at any time and on a school-by-school basis. The mayor of each municipality was responsible for the takeover decision, though the city council had the power to block the mayor's decision. The program was characterized not only for its size in terms of the number of pupils affected (over 5 million) but also for its longevity, since the PSDB maintained the program after winning two succeeding state elections in 1998 and 2002.

During the pre-FUNDEF period (1996 and 1997), there was no pre-established financial compensation in exchange for school adoption. Financial compensations were negotiated on a case-by-case basis, depending on the financial situation of each municipality and on the number of schools they were willing to adopt. However, the FUNDEF reform granted to the adopting municipalities the necessary financial resources to manage the schools, since the fund's resources were allocated in order to maintain the spending per pupil constant. After its implementation, FUNDEF comprised all the fiscal incentives for the adopting municipalities.

Before program implementation, the vast majority of São Paulo's municipalities had no expertise with primary or secondary education provision, being only responsible for kindergarten and preschool administration. Therefore, participation in the program represented a significant administrative challenge for the adopting municipalities as a result of the higher complexity involved in primary and secondary education vis-à-vis the lower levels of education. We thus take this context as particularly prone to have been permeated by an exchange of information among mayors over school adoption procedures and consequences. In particular, we investigate whether the municipalities' school adoption behavior was mainly driven by electoral concerns or by genuine intents of improving the public education provision, i.e. what kind of information was mostly valued and incorporated by mayors in their decision process.

Upon school adoption, the property rights of all school physical resources, including the building itself, are permanently transferred to the municipal government. Some of the school's human resources, such as teachers and staff, are temporarily lent by the state to the municipal administration until the municipalities hire their school professionals to attend the demand of the newly adopted school. According to the laws regulating the transfers of schools, municipalities are fully responsible for all school management activities, from setting the school curricular core to designing school pro-

fessionals' career plans. Furthermore, once a municipality adopts[8] a school, its students are automatically transferred. State legislation mandates that public school students must attend the public school nearest to their homes, irrespective of whether it is state of municipal.

Tables 1 and 2 illustrate the diffusion of the decentralization process from 1996 to 2003. Table 1 displays the evolution of the number of primary schools adopted by the 640 (of 645 São Paulo's) municipalities in the data, and table 2 reports the evolution of enrollment in adopted schools. By 2003, approximately 2,800 primary schools had been adopted, accounting for over 40% of the schools managed by the state administration in the beginning of the reform.

Table 1: Municipalities' adhesion to the school decentralization program

	Year							
Decentralized Schools	1996	1997	1998	1999	2000	2001	2002	2003
0	607	404	351	250	241	214	178	161
1	20	82	101	116	115	122	135	139
2	3	45	56	73	75	75	75	81
3	1	24	27	43	44	45	51	49
4	1	9	12	20	19	23	26	28
5	2	15	19	23	22	23	24	21
6		12	15	24	25	25	26	31
7		10	10	16	15	19	16	16
8		5	7	14	16	16	15	16
9		7	9	9	11	9	10	10
10	2	5	6	13	11	12	14	14
11 - 20	3	14	17	27	33	42	51	55
21 - 30		6	8	9	9	10	14	13
> 30	1	2	2	3	4	5	5	6
# municipalities	640	640	640	640	640	640	640	640

2.3 The Scope for Information Transmission

As mentioned earlier, the decentralization of schooling in Brazil may have presented some difficulties to the municipalities due to their general inexperi-

[8] At this point, the adoption process deserves a little more clarification. We assume that the municipal management of adopted schools takes place only in the year following the mayor's decision for adoption (i.e. the filing of the adoption request). For instance, if a mayor decided to adopt a school in July 2001, the municipal administration is assumed to have begun its management over the school only in the beginning of the following school year, February 2002. As the interest in this analysis relies on the officials' decision-making over decentralization, "adoption" and "decentralization" will henceforth be used to refer to the event of the respective decision.

Table 2: Evolution of adopted schools and students

	# adopted[9]	
Year	Schools	Students
1996	137	
1997	914	36104
1998	236	234472
1999	613	70120
2000	130	176065
2001	247	39053
2002	294	49040
2003	145	78244
Total	2716	683098

ence in providing primary and secondary education services. However, since 1986 Brazilian municipalities have counted on an interesting tool for aiding and shaping their educational policy making, made possible by the creation of the *União Nacional dos Dirigentes Municipais de Educação* (UNDIME). The UNDIME was originally conceived as an association of municipal counselors of education from the state of Pernambuco, who expressed the common desire to democratize the public provision of education through decentralization. This initiative received political support from the Ministry of Education and soon achieved national consolidation. It explicitly stated as its primary objective the need to further insert such issues in the national educational debate, to bring this debate to the municipal level and to uphold the interests of municipal administrators in the educational policy decision process.

In particular, its operation involved the promotion of training activities for the municipal counselors of education, as well as the establishment of regional networks for the exchange of information and experiences among such councils in conferences related to public education. Figure 1 below depicts the configuration of São Paulo municipalities into the 46 groupings within which these events were organized. Moreover, the UNDIME keeps in touch with unions and other civil associations, encouraging the participation of different levels of society in the educational process.

Hence, it seems that the UNDIME may have performed a fundamental role in the process of decentralization of schooling by lowering the learning costs faced by mayors (and counselors) over the provision of primary education. As described in the next section, our empirical analysis of the political determinants of the decentralization process in the state of São Paulo is fundamentally based on such potential influence.

Figure 1: São Paulo Municipalities and UNDIME groupings

3 Data

This section describes our data and the construction of our measures of information flow. The information used in this paper comes from three different sources of Brazilian data: school administrative data, Decennial Population Census data, and political data regarding state and municipal elections.[10]

3.1 Main Data Sources

School data is mainly taken from the School Census, an annual survey that collects information on every school in Brazil, both public and private. The survey is conducted by the Ministry of Education in collaboration with state-level education departments. Questionnaires are sent to each school principal and response is mandatory. The data provides detailed information on school physical resources, such as number of classrooms, computer labs and access to water and sewerage, and human resources such as number of teachers per school level and the highest degree of education obtained by the teachers. At the student level, it also provides information on the number of students per school grade and on student performance in the form of failure and dropouts per school grade. We also have information, provided by the

[10]School data corresponding to the years 1996-1998 appear to be of very poor quality and we refrain from using it in our analysis. We also discard the use of fiscal data since most of the main fiscal variables at the municipal level had relevant changes in their definition after 2001.

São Paulo State Council of Education, regarding the timing of the adoption of state schools by the municipalities.

The Decennial Population Census is the most detailed Brazilian household survey. It has been collected decennially since 1950 by the Brazilian Institute of Geography and Statistics (IBGE), an agency of the federal government.[11] The census data is organized into two different samples, the sample census and the universal census. The former provides information on the universe of the Brazilian households, while the later provides more detailed household information for a sample of 20% of the Brazilian households (although limited to the head-of-household). This paper uses data from the census tracts, which are constructed based on the universal census, for the years 1991 and 2003. The census tracts are geographic divisions of the census containing information on roughly 1000 households each.[12] The variables available provide information on several housing characteristics (e.g. access to treated water and sewerage, number of bathrooms) and on household members, such as age, gender and literacy (income and years of education are also available, although only for the head-of-household.

From 1996 to 2003, four elections were held in Brazil, one every two years starting from 1996. In the years 1996 and 2000, local elections were held for mayors and city council, and in 1998 and 2002 general elections were held for president, state governor, the senate, the state congress and the national congress. The electoral data provides information on all election outcomes per polling station, such as the number of voters, turnout rate and the proportion of votes received by all candidates and political parties.

For the urban municipalities with more than 25,000 habitants in 2000 (which accounts for 170 municipalities out of 645), the IBGE provides digital maps of the census tracts, and the SEADE (Data Analysis Foundation of the State of São Paulo) provides digital street maps. By combining these maps, it becomes possible to identify in which census tract each school is located through the full school addresses provided by the school census. Making use of GIS techniques and interpolating the 1991 and 2000 Population Census, we have aggregated the census tracts for each public school neighborhood by year, where the school neighborhood was defined to match the area comprising its potential users, accordingly to the Brazilian legislation. The data appendix explains the construction of the school neighborhoods in further detail.

Our final data set is a five-year school level panel (from 1999 to 2003) including information on 1598 public schools located in the state of São Paulo

[11] The only exception was in the 90's, where the census was collected in 1991.

[12] Due to the IBGE confidentiality policy, census tract micro data are not available; IBGE only provides the marginal distribution of each variable. Census tracts' borders are defined by the IBGE according to administrative criteria. The 2000 census data for the state of São Paulo was organized into 49,713 census tracts (in 645 municipalities).

and managed by the state administration in 1999.[13] Besides the school level information available from the school census data, the panel also contains yearly information on the census tracts household variables for the public school neighborhood. Furthermore, we associate to each school the electoral data corresponding to the polling station closest to it (whenever the school itself is not employed as a polling station).

3.2 Information Transmission among Municipalities

Our approach assumes that the mayors' decision making process over school decentralization encompasses gathering information not only about the bureaucratic procedures involved but also, and more importantly, about the impacts of adoption on distinct dimensions of their subjective preferences. More precisely, we consider that each adoption experience conveys information in two key aspects: the effect of decentralization on the quality of education provided in the school (i.e. the potential by municipal administrations of performing better than the state administration) and its impact on the mayor's political acceptance (and electoral performance) in the school neighborhood.[14]

Furthermore, we admit that after finding out the consequences of their private adoption experiments, municipal administrations exchange these two kinds of information among themselves. Our primary interest thus lies on determining the role of social learning on the evolution of the schooling decentralization process in São Paulo. That is, we investigate how mayors' decisions respond to the actions and outcomes of other mayors in their information networks and, in particular, which dimension tends to be of higher concern.

In order to disentangle the effects of information transmission within an UNDIME grouping from spatially correlated unobservable shocks (or even other kinds of geographically-based information exchange) affecting the municipalities' adoption decisions, we allow for some heterogeneity in the ability to learn from others. In this sense, for each municipality m we consider three distinct information neighborhoods: (i) the one constituted by contiguous municipalities in its UNDIME grouping (GC_m); (ii) the municipalities in its UNDIME grouping but non-contiguous (GNC_m); and (iii) contiguous municipalities belonging to other UNDIME groupings (NCG_m).[15] An example

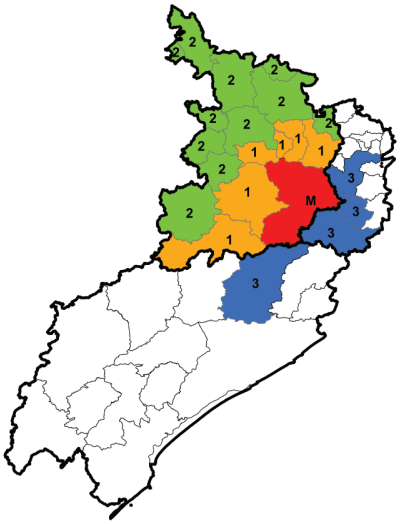
[13] As described in the next section, we exclude those schools located in municipalities presenting at least one empty information neighborhood.

[14] Naturally, in order to treat such effects as independent, we must assume that voters have different (either superior or inferior) information or distinct preferences regarding education provision relative to the government officials.

[15] For instance, we might think that the information conveyed in UNDIME groupings meetings may contain details whose revelation is not permitted by mere geographic proximity. On the other hand, in a grouping meeting, a counselor may value the information conveyed by a geographically close colleague differently from that transmitted by a col-

of such construction is displayed in figure 2 below for the municipality of Ibiuna (in red).[16]

Figure 2: Information Neighborhoods for the Municipality of Ibiuna



We now discuss the construction of our measures of information flow among municipalities. First, for each (adopted) school district, we consider the variation in the performance of the adopting mayor's party between the election preceding the adoption and the election right after.[17] A decentralized school is said to be a *political success*, in a given year, if two conditions hold: (*i*) the decision for its adoption was made between the two latest elections, and (*ii*) the mayor's party performance in the district was below its municipal mean in the election before the adoption and became above the mean in the succeeding election.[18] Similarly, an adopted school constitutes a *political failure* in a given year if (*i*) holds and if the mayor's party performance in the district was above its municipal mean in the election before the adoption and became below the mean in the succeeding election.

league from a more distant municipality.

[16]In particular, those schools located in municipalities in the interior (i.e. not in the border) of a UNDIME grouping (roughly 25% of the municipalities) are excluded from the sample, since they do not have the corresponding type (*iii*) neighborhood.

[17]For general elections, a party performance in the district is taken to be that of its national congress candidates. For local elections, it is simply the performance of its candidate for mayor.

[18]In particular, we assume that a decentralized school no longer contributes with information on the political dimension (for the mayor in exercise) after two years of its adoption.

Note that a school may be seen as a political success/failure by a mayor not responsible by its adoption (in case the adopting mayor is not reelected).[19]

Second, we look at the associated variation in the quality of education provision, proxied by the school's average dropout rate.[20] An adopted school is then said to be a *quality success* in year t if its dropout rate was above the municipal mean in the year preceding decentralization and below the mean in year $t-1$. Schools constituting a *quality failure* are defined accordingly.[21]

Finally, for each municipality m we construct the difference between the number of political successes and failures and the difference between quality successes and failures (henceforth called political and quality deltas and notated $politicaldelta_m$, $qualitydelta_m$) faced by the corresponding administration each year. We assume that these differences capture all the relevant information concerning the decentralization process possibly transmitted among municipalities. Thus, for a given municipality m, we employ the sum of the political (alternatively, quality) deltas faced by the municipalities in an informational neighborhood of m as a measure of the political (alt., quality) learning by the administration in m stemming from the information flow in the corresponding neighborhood. That is, our six regressors of interest are defined by

$$politicaldelta_{N_m} \equiv \sum_{n \in N_m} politicaldelta_n$$

and

$$qualitydelta_{N_m} \equiv \sum_{n \in N_m} qualitydelta_n,$$

for $N \in \{GC, GNC, NGC\}$.

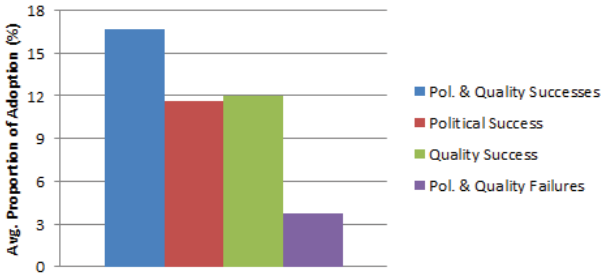
Figure 3 below illustrates the school adoption propensity of municipal administrations without previous adopting events according to their neighbors' experiences (pooling the three neighborhoods). Municipalities are distinguished in four (time-varying) groups: (i) the ones whose neighbors face a positive aggregate political delta; (ii) the ones whose neighbors face a positive aggregate quality delta; (iii) municipalities whose neighbors face positive aggregate deltas in both dimensions; and (iv) municipalities whose neighbors face negative aggregate deltas in both dimensions. Then, for each group, we compute (pooling across years) the average of the proportions of recently adopted schools in the corresponding municipalities.

[19] In this case the comparison made by the mayor concerns the performance of its predecessor's party.

[20] As a robustness check, we alternatively consider using each school's average failure rate as a measure of the quality of education. However, the dropout rate seems to be a better measure since schools may differ on their promotion policies.

[21] Note that for the quality measures employed, higher values are associated with a lower quality of education.

Figure 3: Propensity for School Adoption



There we observe that municipalities bearing witness to more successful adopting experiences (in whichever dimension) by their neighbors appear to be more inclined to begin the implementation of the decentralization reform. Although preliminary, this provides some evidence in favor of our hypothesis of social learning.

4 Empirical Strategy

Our primary method to test for social learning is to estimate a model of the probability that a given school i is under municipal administration (decentralized) in year t, conditional on not having been so before.[22] The focus on this element is due to three reasons. First, it is easily interpreted as the fundamental policy instrument available to elected officials at each time, namely (and roughly) the propensity to adopt a school among those (still) at hand. Second, it constitutes a more basic quantity in the sense that other variables of potential interest (e.g. the probability that the decision for the adoption of i is made in a given year) may be computed from it, but the opposite does not necessarily hold. Finally, it allows for more flexible modelling by implicitly letting such derived quantities depend on regressor vectors of different dimensions in distinct time periods.

Besides, we find the use of a linear model for the conditional probability to be inadequate. One would expect school features accounting for a lower probability of adoption to be more and more concentrated in non-decentralized schools as time goes by. Thus, in light of this sorting effect,

[22]We assume that once a school is decentralized, it cannot ever be returned to the state administration, a hypothesis pretty much backed up by the data. This means that decentralization in t would completely determine the probability of being under municipal administration after t. In practice, this amounts to disregarding from the regressions the observations for schools decentralized for more than one year, since they do not contribute in explaining the adoption decision.

pooling observations across years (even controlling for time effects) would probably yield inconsistent estimates. Rather, we treat this conditional probability as the corresponding hazard function of a duration model for the time until decentralization.[23] Formally, let T_i be the time elapsed until the adoption of school i.[24] Also, let $\mathbf{x}(t)$ denote the value of a vector of (possibly time-varying) covariates in t and let $\mathbf{X}(t)$ be the path of $\mathbf{x}(s)$ from 0 to t.[25] The probability that school i is adopted in year t_j given that it was still under state administration in year $t_j - 1$ (and conditional on the covariate path) is then expressed by

$$\theta(t_j; \mathbf{X}(t_j)) = P(T = t_j \,|\, T \geq t_j, \mathbf{X}(t_j)).$$

Despite the impossibility of taking these expressions as conditional probability density functions (since its conditioning events, $\mathbf{X}(t)$, depend on its argument, t), they can still be employed in the construction of a partial likelihood, which will provide our estimates. In order to do so, however, we must assume that the regressor variables are constant within years (i.e. the time intervals in which the observed durations are known to be). The corresponding log-likelihood function is given by

$$\log L = \sum_{i=1}^{N} \sum_{t_j=0}^{t_i} y_{ij} \log \theta(t_j; \mathbf{X}_i(t_j)) + (1 - y_{ij}) \log[1 - \theta(t_j; \mathbf{X}_i(t_j))]$$

where t_i stands either for the year of school i's adoption or for the year when it is last observed (in case of censoring) and $y_{ij} = 1$ if $j = t_i$ and 0 otherwise.[26]

Following standard practice in the empirical survival analysis literature, we parameterize the hazard function using a proportional hazards form[27], i.e.

$$\theta_i(t; \mathbf{X}(t)) = \theta_0(t) \exp\{\mathbf{x}_i(t)'\boldsymbol{\beta}\} \tag{1}$$

We estimate the baseline hazard, $\theta_0(t)$, nonparametrically through the use of year dummies. As discussed in Meyer (1990), this semiparametric approach may incur in a small efficiency loss but insures consistency, while a parametric assumption for the baseline hazard inconsistently estimates $\boldsymbol{\beta}$

[23] A duration model also provides a natural way to deal with the censoring in a large proportion of such spells and can be more flexible about the shape of the distribution than Tobit-type techniques.

[24] We consider the individual durations as i.i.d. realizations of a random variable T

[25] In this setting, $t = 0$ corresponds to year 1996, the beginning of the decentralization program.

[26] Although stating the likelihood as the one derived from a discrete duration model, our model may be interpreted as an incompletely observed continuous duration model. For further details, see Lancaster (1990) and Jenkins (1995).

[27] In particular, we assume that our choice of the regressor vector is such that $\mathbf{x}(t)$ captures all relevant information in $\mathbf{X}(t)$ regarding the behavior of the hazard function.

when the assumed baseline is incorrect. Also, we allow for the presence of multiplicative unobserved heterogeneity. The (conditional) hazard then becomes

$$\theta_i(t; \mathbf{X}(t)) = \nu_i \theta_0(t) \exp\{\mathbf{x}_i(t)'\boldsymbol{\beta}\} \tag{2}$$

where ν_i is a random variable assumed to be independent of $\mathbf{x}(t)$ and with mean one. The corresponding log-likelihood is obtained by conditioning on ν_i and integrating over its distribution.[28] As shown in Lancaster (1979) for the case of time-invariant regressors, not allowing for such heterogeneity (in case it really exists) induces an attenuation bias in the estimates of $\boldsymbol{\beta}$. Also, we may interpret ν as representing the effect of measurement error in the covariates and/or in the duration T as well as the omission of time-invariant explanatory variables.

5 Empirical Results

The base regression specification that we use to investigate the relationships between the probability that an available school be adopted and information shocks is defined by setting

$$\begin{aligned}\mathbf{x}_i(t)'\boldsymbol{\beta} = {} & \alpha^{GC}_{pol} politicaldelta_{GC_{m_i},t} + \alpha^{GC}_{qual} qualitydelta_{GC_{m_i},t} + \\ & + \alpha^{GNC}_{pol} politicaldelta_{GNC_{m_i},t} + \alpha^{GNC}_{qual} qualitydelta_{GNC_{m_i},t} + \\ & + \alpha^{NGC}_{pol} politicaldelta_{NGC_{m_i},t} + \alpha^{NGC}_{qual} qualitydelta_{NGC_{m_i},t} + \\ & + \mathbf{z}'_{i,t}\boldsymbol{\gamma}\end{aligned} \tag{3}$$

where m_i represents school i's municipality and $\mathbf{z}_{i,t}$ denotes a vector of conditioning variables. These include school i's characteristics as well as political and demographic information concerning both school i's district and municipality, and UNDIME grouping fixed effects. All specifications also include as controls the sizes of municipality m_i's neighborhoods, the number of schools adopted in each of its neighborhoods until the year before (excluding its own adoption experiments), and the number of state (i.e. available) schools in m_i in 1996. We refrain from including a municipality's own adoption experience so far in order to avoid endogeneity issues. Nevertheless, it seems reasonable to assume that the neighborhoods are large enough to make the information arising from one's own experience negligible in comparison to that generated by the neighbors' experiences. Alternative specifications and robustness checks are discussed in the next sections.

Our identification assumption is thus that conditional on such characteristics of the school's political and economic environment mentioned above,

[28] For computational convenience, we assume the ν_i's to be independent and lognormally distributed.

our measures of information flow are uncorrelated to unobserved determinants of the choice for school adoption. In other words, successful adoption events experienced by a given mayor's information neighbors influence his decision over the adhesion to the decentralization program only through the suggested information links. Therefore, if social learning over the decentralization process' impacts is made possible by an information flow within a certain kind of neighborhood N, we would expect α^N_{pol} and/or α^N_{qual} to be positive (depending on which kind of information is transmitted and valued by the agents).

Table 3 presents our main results, which are based on model (1). Columns (1) and (2) firstly show estimates for the effect of social learning considering each of the dimensions of potential interest to the mayors separately, while column (3) displays the estimates from specification (3) (controlling for both sets of measures). Estimating model (2) with these specifications (and those in the next section) yields extremely similar results and are therefore suppressed.[29] We see that a larger number of politically successful experiences by a mayor's UNDIME grouping peers (either contiguous or not) is associated with a higher probability that he will choose to adopt the schools in his jurisdiction. The two corresponding estimated coefficients are positive, statistically significant, and large (specially the one for contiguous peers).

Given the form of the hazard function defined in (1), we have $\frac{\partial \ln \theta}{\partial x_j} = \beta_j$. Hence, an increase of one (net) political success within a mayor's UNDIME-grouping contiguous neighbors is associated with an increase by roughly 30% of his initial probability of adopting a given school (from 20% to 26%, for instance). Thus, political learning seems to have played an important role in the evolution of the decentralization process. On the other hand, we find no statistically significant evidence for the effects of education-related information exchange (and the estimates for such flow regarding UNDIME-grouping peers become lower after we include the political measures in column (3). In particular, it is possible that the meetings promoted by the UNDIME may have subverted the main focus of its debate (although it is not yet clear how or whether such higher relevance of political concerns adversely affects the public education provision). That is, the participating education counselors seem to have been more concerned about the political dimension of the returns to school adoption, either in sorting the details to be communicated or in absorbing them.

[29] The parameter estimates are identical to those obtained from model (1), and the standard errors are only slightly different (in such a way that no conclusions concerning statistical significance are altered).

Table 3: Conditional Probability of Decentralization - Discrete-Time Proportional Hazards Model

	(1)	(2)	(3)
$politicaldelta_{GC_m}$	0.311^{***}	-	0.298^{***}
	(0.107)		(0.115)
$politicaldelta_{GNC_m}$	0.142^{**}	-	0.153^{***}
	(0.0582)		(0.0562)
$politicaldelta_{NGC_m}$	-0.0775	-	-0.0739
	(0.0807)		(0.0885)
$qualitydelta_{GC_m}$	-	0.251	0.131
		(0.215)	(0.219)
$qualitydelta_{GNC_m}$	-	0.195	0.115
		(0.181)	(0.174)
$qualitydelta_{NGC_m}$	-	0.287	0.330
		(0.289)	(0.297)
Observations	6,273	6,273	6,273
Schools	1,598	1,598	1,598
Log-likelihood	-352.2	-361.7	-350.7

*, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively. Standard errors clustered at municipality level. Controls include electoral data, school resources, demographic characteristics of the school district and municipality, year dummies and UNDIME-grouping dummies.

6 Robustness Checks and Extensions

6.1 Alternative Quality Measures, Information Neighborhoods and Confounding Factors

The lack of relevance of education quality implications of school decentralization suggested by our results may simply be driven by the fact that the school's dropout rate may not be seen by mayors as an appropriate quality measure. We thus consider defining alternative quality successes/failures based on the evolution of the schools' failure rate. The results obtained reestimating our model with these measures are presented in Table 4. We see that, despite the fact that the schools' dropout and failure rates are poorly correlated[30] in our sample, the results are virtually unchanged.

On the other hand, it is possible that the political success of a mayor's party is correlated to quality improvements on the provision of public education that are not captured by such measures. We explore this issue through the use of data on the students' performances in a federal standardized exam

[30]The correlation of such variables in our sample is 0.10.

Table 4: Conditional Probability of Decentralization - Discrete-Time Proportional Hazards Model

	(1)	(2)	(3)
$politicaldelta_{GC_m}$	0.311^{***}		0.316^{***}
	(0.107)		(0.115)
$politicaldelta_{GNC_m}$	0.142^{**}		0.158^{***}
	(0.0582)		(0.0616)
$politicaldelta_{NGC_m}$	-0.0775		-0.0802
	(0.0807)		(0.0803)
$qualitydelta_{GC_m}$		-0.0039	-0.0743
		(0.0902)	(0.0953)
$qualitydelta_{GNC_m}$		0.107	0.0739
		(0.120)	(0.154)
$qualitydelta_{NGC_m}$		0.0448	-0.0239
		(0.120)	(0.130)
Observations	6,273	6,273	6,273
Schools	1,598	1,598	1,598
Log-likelihood	-352.2	-361.7	-350.7

See notes to Table 3.

(SAEB), applied in a randomized sample of schools. As shown in Table 5, the regression of any of the proficiency measures (in either Portuguese or math tests) against the corresponding school-year dropout rate and the performance of the mayor's party in the preceding election suggests that our measures of political success are not correlated with other school quality measures (i.e. the estimated coefficients for political performance are statistically insignificant).

A more serious matter is the possibility that our measures of information flow be correlated with unobserved determinants of school decentralization. We argue that this is not so. Firstly, we present some evidence that our information measures are uncorrelated with most of the unobserved determinants of adoption. This can be seen in Tables 6 and 7, where our school quality measures are based on the dropout rate and on the failure rate, respectively.

We observe that dropping our control variables from the estimated model has very little effect in our conclusions, except when UNDIME-grouping fixed effects are not accounted for, which turns our political-grouping estimates insignificant. In this sense, we argue it is plausible that any existing (unobserved) confounding factors distorting our results are most likely grouping (spatially)-based.

Table 5: Correlation of Political and Quality Success Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Math		Portuguese		Math		Portuguese	
Dropout Rate	−1.405	−1.840	−2.651**	−3.096**	−2.584*	−2.652*	−3.702**	−3.920***
	(1.279)	(1.301)	(1.264)	(1.290)	(1.424)	(1.447)	(1.419)	(1.427)
Past Pol. Performance	0.163	0.166	0.00929	−0.0184	-	-	-	-
	(0.103)	(0.141)	(0.101)	(0.140)				
Future Pol. Performance	-	-	-	-	−0.00311	0.0681	0.00294	−0.207
					(0.116)	(0.215)	(0.115)	(0.212)
Constant	182.3***	189.7***	180.6***	187.5***	191.5***	195.4***	184.3***	196.9***
	(3.434)	(5.092)	(3.393)	(5.052)	(5.029)	(11.97)	(5.014)	(11.80)
Year Dummies	-	Yes	-	Yes	-	Yes	-	Yes
Observations	117	117	117	117	61	61	61	61
R-squared	0.031	0.067	0.037	0.066	0.054	0.056	0.106	0.127

Table 6: Conditional Probability of Decentralization - Discrete-Time Proportional Hazards Model

	(1)	(2)	(3)	(4)	(5)	(6)
$politicaldelta_{GC_m}$	0.298^{***}	0.323^{***}	0.327^{**}	0.220	0.221	0.163
	(0.115)	(0.124)	(0.146)	(0.169)	(0.183)	(0.149)
$politicaldelta_{GNC_m}$	0.153^{***}	0.151^{***}	0.141^{**}	0.0558	0.0372	0.0381
	(0.0562)	(0.0561)	(0.0549)	(0.0516)	(0.0485)	(0.0392)
$politicaldelta_{NGC_m}$	−0.0739	−0.0817	−0.0835	−0.0600	−0.0565	-0.131^{*}
	(0.0885)	(0.0889)	(0.105)	(0.0797)	(0.0766)	(0.0745)
$qualitydelta_{GC_m}$	0.131	0.0918	0.136	−0.0992	−0.0939	−0.158
	(0.219)	(0.222)	(0.235)	(0.214)	(0.216)	(0.197)
$qualitydelta_{GNC_m}$	0.115	0.0799	0.0831	0.0211	0.0327	0.156
	(0.174)	(0.179)	(0.187)	(0.150)	(0.153)	(0.152)
$qualitydelta_{NGC_m}$	0.330	0.262	0.241	0.155	0.170	0.522
	(0.297)	(0.294)	(0.301)	(0.321)	(0.320)	(0.393)
Political	Yes	-	-	-	-	-
School Resources	Yes	Yes	-	-	-	-
UNDIME-grouping Fixed Effects	Yes	Yes	Yes	-	-	-
Demographics (School level)	Yes	Yes	Yes	Yes	-	-
Demographics (Municipality level)	Yes	Yes	Yes	Yes	Yes	-
Observations	6,273	6,273	6,273	6,273	6,273	6,273
Schools	1,598	1,598	1,598	1,598	1,598	1,598
Log-likelihood	−350.7	−358.3	−385.1	−435.1	−442.7	−467.4

See notes to Table 3.

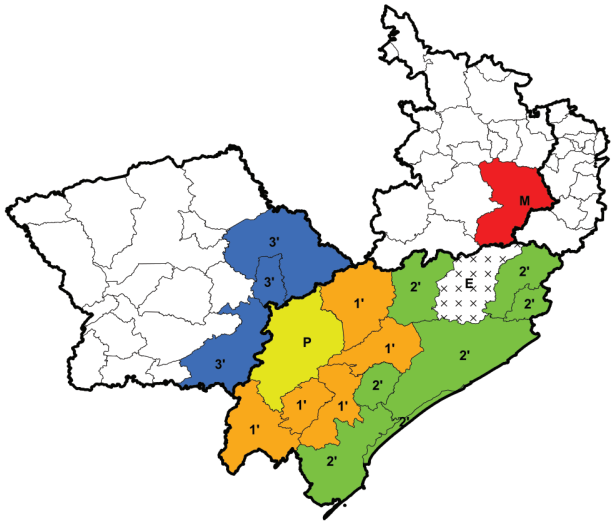
Table 7: Conditional Probability of Decentralization - Discrete-Time Proportional Hazards Model

	(1)	(2)	(3)	(4)	(5)	(6)
$politicaldelta_{GC_m}$	0.316^{***}	0.333^{***}	0.338^{**}	0.203	0.204	0.165
	(0.115)	(0.118)	(0.139)	(0.150)	(0.158)	(0.113)
$politicaldelta_{GNC_m}$	0.158^{**}	0.153^{***}	0.146^{**}	0.0408	0.0223	−0.0139
	(0.0616)	(0.0574)	(0.0580)	(0.0411)	(0.0395)	(0.0473)
$politicaldelta_{NGC_m}$	−0.0802	−0.0884	−0.0877	−0.0336	−0.0272	−0.106
	(0.0803)	(0.0801)	(0.0965)	(0.0712)	(0.0674)	(0.0876)
$qualitydelta_{GC_m}$	−0.0743	−0.0290	−0.0565	0.0542	0.0407	0.113^{*}
	(0.0953)	(0.0860)	(0.0825)	(0.0749)	(0.0765)	(0.0655)
$qualitydelta_{GNC_m}$	0.0739	0.129	0.176	0.220^{*}	0.219^{**}	0.188^{**}
	(0.154)	(0.166)	(0.171)	(0.113)	(0.109)	(0.0958)
$qualitydelta_{NGC_m}$	−0.0239	−0.0467	0.0145	0.0675	0.0727	0.0139
	(0.130)	(0.128)	(0.189)	(0.137)	(0.110)	(0.152)
Political	Yes	-	-	-	-	-
School Resources	Yes	Yes	-	-	-	-
UNDIME-grouping Fixed Effects	Yes	Yes	Yes	-	-	-
Demographics (School level)	Yes	Yes	Yes	Yes	-	-
Demographics (Municipality level)	Yes	Yes	Yes	Yes	Yes	-
Observations	6,273	6,273	6,273	6,273	6,273	6,273
Schools	1,598	1,598	1,598	1,598	1,598	1,598
Log-likelihood	−351.6	−358.6	−384.7	−431.9	−439.5	−468.2

See notes to Table 3.

Secondly, in order to further explore such circumstance, we conduct the following exercise: (i) for each municipality m, we randomly select a municipality n from those belonging to an UNDIME-grouping contiguous to that containing m. We then construct new "synthetic" neighborhoods for m consisting of the corresponding n-neighborhoods[31] discounted by those municipalities originally belonging to any of the three m-neighborhoods. This procedure is illustrated in Figure 4 below for the municipality of Ibiuna (in red), where the selected municipality (Eldorado) is displayed in yellow and the discarded neighbors are crossed. (ii) Then, we recompute our measures of information flow based on these artificial neighborhoods and reestimate model (1) using such measures. Steps (i) and (ii) are replicated 1,000 times.[32]

Figure 4: Illustration of the Reconfiguration of Neighborhoods



Finally, for each of our six regressors of interest, we nonparametrically estimate the probability distribution of the corresponding parameter estimators (obtained by redefining the neighborhoods to preserve some spatially correlated factors, but abandoning the UNDIME-grouping structure). We

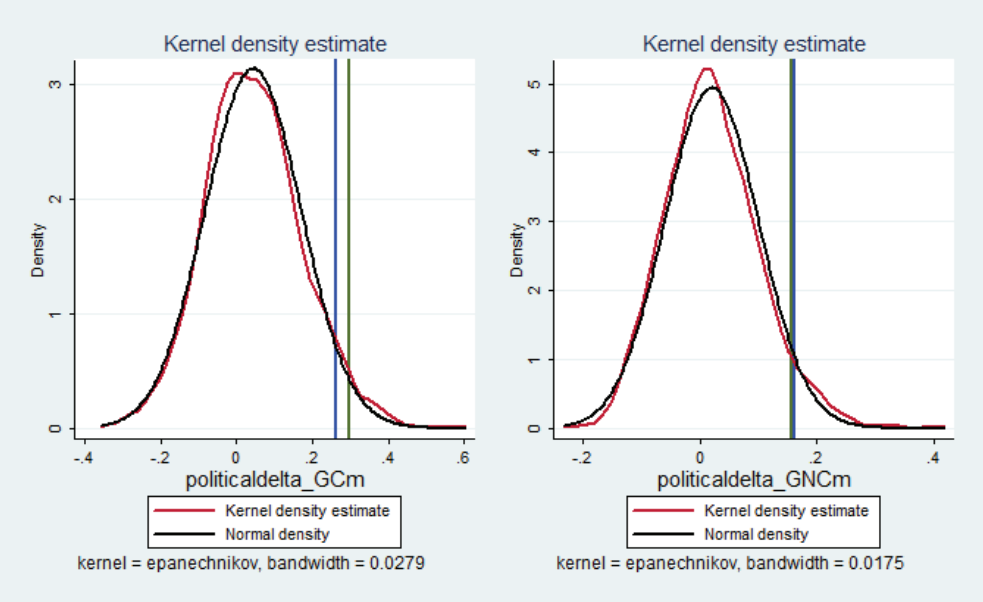
[31] That is, m's new neighborhood of contiguous UNDIME-grouping municipalities neighborhood is taken to be n's contiguous UNDIME-grouping municipalities, and so on.

[32] This discounting of municipalities may in principle lead to the construction of empty neighborhoods for some municipalities. Thus, in order to obtain the same (original) estimating sample in every replication, the randomization in step (i) is remade until no municipalities present empty synthetic neighborhoods.

observe in all cases that the densities are approximately normal and symmetric around zero. Figure 5 below displays the kernel density estimates - and corresponding normal densities - associated with the (artificial counterparts of the) two measures for which a statistically significant effect was found in our main results (column (3) of Table 3), i.e. $politicaldelta_{GC_m}$ and $politicaldelta_{GNC_m}$.

In each panel, we also indicate the original estimate of the corresponding effect with a green vertical line and the 95th percentile of the empirical distribution with a blue vertical line. Thus, we observe that our original estimate for the effect of learning from UNDIME-grouping contiguous municipalities is above the 95th percentile of the corresponding kernel density estimate, while the original estimate for the effect of learning from UNDIME-grouping non-contiguous municipalities is (above the 90th percentile and) slightly below the 95th percentile of the corresponding density; the four remaining original estimates are between the 10th and 90th percentiles of the corresponding estimated densities. Hence, we obtain some suggestion that other geographically-based factors affecting the evolution of the decentralization program are not driving our main results nor compromising our social learning interpretation.

Figure 5: Robustness Check - Empirical Distributions



6.2 Political Affiliation Heterogeneity in the Extent of Social Learning

The evidence presented in our analysis that electoral considerations may have been the essence of the learning process leading to the development of the decentralization program suggests the question of how effective the information links were, according to the political proximity of neighbors. To investigate this matter we compute, for each municipality m, the proportion (assuming values from zero to one) of municipalities, in each of its neighborhoods, administered by mayors affiliated to the same party as m's ($\%sameparty_{N_m}$, for $N \in \{GC, GNC, NGC\}$). We then construct interactions of each of our measures of information flow with corresponding neighborhood proportion of party "colleagues". Table 8 presents the results of reestimating model (1) adding those interactions to specification (3). When the dropout rate is used to construct our quality measures (Column (1)) we find evidence that political affiliation may have affected the degree to which communication about school adoption was effective (one might think that officials affiliated to opposing parties may have reasons not to trust each other's advices), at least in the neighborhood of non-contiguous UNDIME-grouping peers. The coefficients on the interactions corresponding to that neighborhood are positive and statistically significant.

Our results suggest for instance that, all else equal, a mayor with 10% more party colleagues in that neighborhood, when faced with a net increase of one political success (in that same neighborhood) is associated with an adoption probability 35% higher.[33]

The reason why a similar pattern is not found for the neighborhoods of contiguous municipalities may be due to the possibility that geographic proximity enables (and induces[34]) mayors to corroborate by themselves any information transmitted by their peers. In this sense, mayors would not distinguish the information received from contiguous neighbors according to their parties. On the other hand, when our quality measures are based on the school's failure rate, no statistically significant effect is found for the introduced interactions.

7 Conclusion

This paper analyzes a specific scenario of decentralization of political responsibilities over the public provision of services - that of primary schooling in the state of São Paulo - in an attempt to investigate the relative relevance

[33] In fact, if we interpret one such interaction as the expected number of net successes (in the respective neighborhood) communicated by party colleagues, our results suggest that the impact of a net success transmitted by a party colleague is more than 20 times that of a net success transmitted by officials affiliated to other parties.

[34] Via *yardstick competition*, for instance.

Table 8: Conditional Probability of Decentralization - Discrete-Time Proportional Hazards Model

	(1)	(2)
$politicaldelta_{GC_m}$	0.325^{***}	0.328^{***}
	(0.113)	(0.102)
$politicaldelta_{GC_m} * \%sameparty_{GC_m}$	-0.511	0.631
	(1.171)	(1.559)
$politicaldelta_{GNC_m}$	0.151^{***}	0.167^{***}
	(0.0574)	(0.0624)
$politicaldelta_{GNC_m} * \%sameparty_{GNC_m}$	3.516^{**}	1.191
	(1.613)	(0.989)
$politicaldelta_{NGC_m}$	-0.0778	-0.0716
	(0.0894)	(0.0835)
$politicaldelta_{NGC_m} * \%sameparty_{NGC_m}$	-1.223	-1.784
	(1.820)	(2.223)
$qualitydelta_{GC_m}$	0.208	-0.0843
	(0.232)	(0.0940)
$qualitydelta_{GC_m} * \%sameparty_{GC_m}$	0.373	-0.376
	(0.749)	(1.428)
$qualitydelta_{GNC_m}$	0.0543	0.0943
	(0.167)	(0.186)
$qualitydelta_{GNC_m} * \%sameparty_{GNC_m}$	3.076^{*}	-3.147
	(1.742)	(3.475)
$qualitydelta_{NGC_m}$	0.315	-0.0332
	(0.269)	(0.129)
$qualitydelta_{NGC_m} * \%sameparty_{NGC_m}$	$-^{a}$	-2.321
		(5.080)
Observations	6,273	6,273
Schools	1,598	1,598
Log-likelihood	-337.4	-349.3

[a]There is insufficient variation in the sample to estimate this coefficient. All other notes to Table 3 apply.

of officials's concurrent motivations over the adhesion to that particular policy, as well as the viability of successful implementation of public policies in general (i.e. how capable are the political interests and present democratic institutions in promoting the alignment of the officials' behavior to the actions intended by policymakers). We argue that such decentralization reforms are characterized by local governments potentially presenting (due to their lack of experience) an inferior information over the necessary procedures for an adequate provision of the service under devolution. Such

occurence, in its turn, could be severely detrimental to the expansion of these reforms whenever local governments may choose to adhere to them, as in our case, and induce the realization of learning efforts. Thus, we focus on the role of social (or, in this case, political) learning on the development of the decentralization reform, exploring how the information exchange among officials is configured and especially which aspects (dimensions of the returns) of the newly assumed responsibilities are mostly valued by them in the learning process.

We present evidence that social learning is indeed an important factor in the diffusion of knowledge regarding primary education provision and in the adhesion to the decentralization program. The challenge of identifying learning effects is adressed by making use of data regarding exogenously defined information neighborhoods combined with conventional economic information. We find that mayors are more likely to "adopt" state schools upon the reception of good news about the electoral returns of decentralization, but seem to be unaffected by considerations of provision quality improvement. That is, mayors tend to assume the administration of state schools in their municipalities when information neighbors experience an electoral success associated to their previous adoption experiences. On the other hand, experiences by information neighbors showing up to be successful in the quality dimension seem to be ignored in the mayors' decision making process over school decentralization. Our suggestion that the learning process was mainly driven by electoral considerations is further supported by the evidence that a mayor tends to be more responsive to the information transmitted by party colleagues, i.e. neighbors affiliated to the same party as his own.

In particular, we conclude that both the presence of social learning among officials and of democratic mechanisms of public choice configured key elements to the extension of the decentralization reform. In this sense, the promotion of information exchange possibilities should be encouraged and implemented by the higher levels of government in the event of introduction of such programs. On the other hand, the question of whether the adhesion to the program was effectively accompanied by the intended increase in provision quality is a matter left for future research. We also note that, while in this case the information networks were exogenously defined, a proper evaluation of general policies involving the diffusion of knowledge about newly introduced responsibilities/technologies should take into account the potentially endogenous process of information network formation. While our evidence regarding the influence of political affiliation sheds some light on how this process may occur in political contexts, further investigations are required.

References

AGHION, P.; HOWITT, P. Endogenous growth. MIT Press, 1998.

BARDHAN, P. Decentralization of governance and development. *The Journal of Economic Perspectives*, v. 16, n. 4, p. 185–205, 2002.

CONLEY, T. G.; UDRY, C. R. Learning about a new technology: Pineapple in ghana. *The American Economic Review*, p. 35–69, 2010.

EPPLE, D.; ZELENITZ, A. The implications of competition among jurisdictions: does tiebout need politics? *The Journal of Political Economy*, p. 1197–1217, 1981.

FOSTER, A. D.; ROSENZWEIG, M. R. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*, p. 1176–1209, 1995.

GLAESER, E. L. et al. Growth in cities. *Journal of Political Economy*, v. 100, n. 6, p. 1126–52, December 1992.

JENKINS, S. P. Easy estimation methods for discrete-time duration models. *Oxford bulletin of economics and statistics*, v. 57, n. 1, p. 129–136, 1995.

LANCASTER, T. Econometric methods for the duration of unemployment. *Econometrica: Journal of the Econometric Society*, p. 939–956, 1979.

LANCASTER, T. The econometric analysis of transition data. Cambridge university press, 1990.

MEYER, B. D. Unemployment insurance and unemployment spells. *Econometrica*, v. 58, n. 4, p. 757–82, 1990.

OATES, W. E. An essay on fiscal federalism. *Journal of economic literature*, v. 37, n. 3, p. 1120–1149, 1999.

SEABRIGHT, P. Accountability and decentralisation in government: an incomplete contracts model. *European economic review*, v. 40, n. 1, p. 61–89, 1996.

TIEBOUT, C. M. A pure theory of local expenditures. *The journal of political economy*, v. 64, n. 5, p. 416–424, 1956.

8 Appendix - School Neighborhood Data

In order to use GIS techniques to construct the public school neighborhood variables, the 645 municipalities of the state of São Paulo were firstly divided into two groups according to the size of their population. The first group, the "large cities group", contains all the 170 urban municipalities with populations larger than 25,000 habitants in 2000. For these municipalities, the IBGE provides digital maps of the census tracts, which can be combined with digital street maps provided by the SEADE (Data Analysis Foundation of the State of São Paulo) for all municipalities. The second group, the "small cities group", includes all urban municipalities with less than 25,000 habitants and all rural municipalities, for which the digital maps of census tracts are unavailable.

Due to the partial availability of the digital census tracts maps, the definition of the public schools' neighborhoods employed varies according to which municipality group a school belongs. We present the necessary steps to create the public school neighborhood variables for each municipality group separately:

8.1 Large Cities Group

For the schools located in large cities, the data are constructed using the following steps:

Step (1): This consisted of finding the cartographic coordinates (latitude and longitude, or the so-called "geo codes") of each school. For 60% of the schools, this could be accomplished using schools addresses and zip codes provided by the school census data and the digital street maps provided by the SEADE. For the remaining 40% of the schools there were no matches between their address and the digital maps. For the public schools with no match, the geo codes were obtained from hard copy maps kept by the public schools administrators (either the municipal or the state secretaries of education), that indicate the schools' location.

Step (2): The second step consisted of defining the public school neighborhoods and then aggregating the census tract data for the defined neighborhoods. Since the law dictates that public school students must attend the closest school to their homes, the neighborhood for each public school was defined as the area closest to that school than to any other public school (the area where the potential public school pupils are located). Due to public schools attrition in the 1996-2003 period and the fact that the neighborhood boundaries are sensitive to the number of schools within the municipality, the public school neighborhoods were redesigned for every year in the sample. Once the boundaries of the school neighborhoods were defined, all the

variables of the census tracts within a school's neighborhood were aggregated to the neighborhood level.

Step (3): This consisted of interpolating the household variables in the 2000 census with the 1991 census. The major problem in performing this interpolation relies on the fact that the 1991 Population Census is not organized in census tracts. However, since both censuses (2000 and 1991) variables are available at the municipality level, the interpolation at the municipal level is possible. Under the assumption that the time variation of the variables aggregated at the municipal level is a good proxy for the time variation of the variables aggregated at the school neighborhood level, it is thus possible to interpolate the variables at the school neighborhood level. In short, we use the same line obtained for the interpolation of a given variable aggregated at municipal level to interpolate the same variable aggregated at the school neighborhood level. Based on this interpolation procedure we inputted the household variables aggregated at the school neighborhood level for every public school.

8.2 Small Cities Group

As a consequence of the unavailability of the census tracts digital maps for the municipalities in the small cities group, it is not possible to identify the census tracts where the schools are located. It is thus impossible to define the schools' neighborhoods as was done for the schools in the large cities group. To overcome this problem, we firstly classified the schools in each municipality into two groups according to the region (urban or rural) where they are located. Then, we took advantage of the fact that it is also possible to identify the region (urban or rural) where the census tracts are located to aggregate the household variables (provided by the census tract) for the rural and urban areas in each municipality. Lastly, the household variables aggregated for the rural areas were distributed uniformly among the schools located in rural areas, while the household variables aggregated for the urban areas were distributed uniformly among the schools located in urban areas. Using the population census data interpolated (between 1991 and 2003) for the rural and urban areas of each municipality, it was then possible to replicate this procedure for all the available years (1996 to 2003).