POLITICAL SELECTION IN THE SKILLED CITY

Antonio Accetturo[♣]

May 2013

Abstract

This paper studies the impact of human capital on the citizens' ability to select high quality politicians in democratic elections. By using a change in the rules for Italian mayoral elections and a diff-in-diffs estimator, I find that cities endowed with a larger amount of human capital tend to elect higher quality politicians. This result is robust to omitted variables or selection issues and holds for a variety of quality indicators.

JEL: D7, I20, H8.
Keywords: Human capital externalities, political selection, cities.

[♣] Structural Economic Analysis Department – Bank of Italy, Rome. Email: antonio.accetturo@bancaditalia.it

I wish to thank Mino Barone, Matteo Bugamelli, Gustavo Canavire-Bacarreza, Francesco Manaresi, Sauro Mocetti, and Massimo Morelli useful suggestions. The views expressed are the author's own and do not necessarily reflect those of the Bank of Italy. Usual disclaimers apply.

1. Introduction

This paper deals with the benefits of education. It aims at analyzing whether more educated citizens generate a better polity. Studies on the impact of schooling on economic variables date back to more than four decades. Most empirical papers indicate that every extra-year of schooling raises the individual wage by 8-12 per cent. The positive effects of education, however, are not necessarily limited to the individual sphere. Economists have recently argued the existence of positive externalities due to the sheer presence of a more educated pool of citizens on the individual productivity. These are defined as *social* returns to education or human capital externalities. Moretti (2004) extensively reviews this literature.

However, social returns may go beyond the labor market dimension. A pool of more educated individuals may exert a positive effect on political outcomes and the quality of institutions.[1] This is a well-known dimension in Political Science: Almond and Verba (1989) point out education is a crucial determinant of "civic culture" and participation in democratic politics as "the uneducated man […] is a different political actor from the man who achieved a higher level of education".

There are two main reasons why education should influence political outcomes. The first relates to political participation: if education increases the cognitive skills of individuals, their costs to grasp information should be lower and, hence, they should vote more frequently. The second relates to the choices of educated individuals. Education might increase the ability to select more able politicians, since it should raise the capability to understand the issues at stake and to monitor elected politicians.

[1] Human capital is also likely to reduce the probability of an individual to engage criminal activity. Despite the enormous policy implication, little evidence is available. Lochner (1999) finds a negative correlation between high school graduation and crime for the US. Lochner and Moretti (2002) use changes in State compulsory school attendance as an instrument for schooling and find that an additional year of schooling reduces the probability to be incarcerated.

Milligan et al. (2004) analyze the first channel and show that education has a positive effect on the likelihood to vote to presidential elections in US over the period 1948-2000.[2]

The aim of this paper is to provide the first causal investigation of the second channel by analyzing the impact of human capital on political choices. This is done by studying the political selection of Italian mayors over the period 1987-1997 and assessing, in particular, whether municipalities endowed with higher human capital end up in electing politicians with ex-ante better characteristics.

The identification of a causal link that goes from citizens' education to political selection is not an easy task. Human capital is not randomly distributed across cities: it tends to be higher in places with high productivity, good amenities, and better institutions. This implies that better administrated cities are likely to disproportionally attract more educated citizens, thus creating a severe problem of reverse causality (Glaeser and Gottlieb, 2009). Moreover, local politicians are usually selected from a pool of local candidate-citizens (Besley and Coate, 1997) thus creating an omitted variable bias.

In the literature on human capital externalities, a typical way to solve this problem is to exploit some institutional factors that exogenously change the human capital distribution across time or areas. For example, Milligan et al. (2004) exploit the cross-state differences in the adoption of compulsory schooling. The same identification is utilized by Acemoglu and Angrist (1999).

In this paper, I resort to a different identification strategy. I make use of a difference-in-differences (diff-in-diffs) framework that exploits a change in the electoral rules for mayors in 1993. New electoral rules raised the incentives to elect higher quality politicians for all citizens; if more educated individuals care more for higher quality candidates, the rise in the quality should be stronger in cities endowed with higher human capital. The use of city level fixed effect helps in controlling for all other fixed characteristics. This kind of identification, that is unusual in the literature on human capital externalities, is widely used

[2] The effect, however, is limited to the cases in which registration for voting lies with the citizen. When the responsibility to maintain electoral registers rests with local government officials (like in the UK) there are no effects of education on political participation.

in applied international trade for the analysis of the heterogeneous effects of trade liberalizations on firm-level outcomes (see, for example, Verhoogen, 2008).

The identification of a causal parameter rests on the idea that 1993 electoral reform greatly changed the incentives to elect a high-quality politician. In the previous context, mayors were elected by city councils, generally among the available councillors. City council elections were dominated by local sections of national parties, voters had little command on the lists that parties presented and, therefore, on the quality of the elected councillors. As a result, mediation by political parties was such that the voters' decisions was mainly *ideological* since the voters' attention was relatively more on parties' political platforms rather than on the quality of politicians. The 1993 reform established direct voting for mayors along with new powers. As a consequence, voters were induced to more carefully evaluate the quality of the candidates besides the usual assessment of the political platform. The change of electoral rules implied a rise in the overall quality of the mayors, measured as schooling, labor market experience and skill-intensity of the mayor's profession.[3]

Results show that, coeteris paribus, the reform increased my main measure of the quality of mayors (years of schooling) by 6.4%.[4] The effect in the municipality at the 75th percentile of the human capital distribution was 1.2 percentage points stronger than the one in the city at the 25th percentile. This implies that the interquartile range in the human capital distribution accounts by almost one fifth (1.2/6.4 = 0.18) of the total average increase after the reform. This is confirmed by using different definitions for the quality of mayors and in a number of robustness checks that take into account omitted variables bias or sorting of citizens.

The paper is organized as follows. Section 2 presents a literature review on human capital externalities and political selection. Section 3 introduces the institutional framework by describing the 1993 reform. Section 4 describes the empirical design. Section 5 presents the data and some descriptive statistics. Section 6 shows the results. Section 7 concludes.

[3] Veronese (2005) provides a theoretical explanation for this effect.

[4] This effect confirms the empirical evidence found by Veronese (2005) on a smaller set of Italian cities.

2. Literature review

This study is at the crossroad of two distinct streams of literature. The first is related to human capital externalities and their effects on the functioning of democratic systems; the second concerns the issue of political selection.

2.1 Human capital externalities and democracy

The idea that education matters for the functioning of democratic systems is not new. As Friedman (1962) points out: "A stable and democratic society is impossible without a minimum degree of literacy and knowledge on the part of most citizens and without a widespread acceptance of some common set of values. Education can contribute to both. In consequence, the gain from education of a child accrues [...] also to other members of the society". Similarly, Lipset (1960) states that prosperity (a characteristic strictly linked to education) stimulates democracy crediting his idea to Aristotle: "From Aristotle down to the present, men have argued that only in a wealthy society in which relatively few citizens lived in real poverty could a situation exist in which the mass of the population could intelligently participate in politics and could develop the self-restraint necessary to avoid succumbing to the appeals of irresponsible demagogues" (p. 75).

Bourguignon and Verdier (2000) provide a theoretical model in which initial income distribution determines education attainment that, in turn, influences political participation. Glaeser et al. (2007) presents a model on the stability of democracies and the role of education. They assume that benefits from democracy are widespread but small in size, while dictators usually concentrate the rents they extract on a narrower support base. This feature entails the fundamental fragility of democratic institutions: narrow groups have strong incentives to overthrow democracy, while the majority has weaker incentives to defend it. Schooling, however, may attenuate this problem as it increases the individuals' interest for politics by lowering the costs for participation to a democratic system. As a consequence, democracies are more stable in areas with a higher endowment of human capital.

Barro (1999) tests the Lipset/Aristotle hypothesis on a panel of over 100 countries. By using either electoral rights or civil liberties as dependent variable, he finds a positive and

significant effect of GDP per capita and years of schooling and a negative effect of gender gap in education. This implies that, holding per capita income fixed, more educated countries end up in being more democratic. This evidence is confirmed by Glaeser et al. (2004), where the probability to switch to a democratic regime is positively associated with the average number of years of schooling of the population. Papaioannu and Siourounis (2008) contribute to this evidence by showing that economic development and, especially, education were the main determinants for transition to stable democracy in the period 1960-2005. Milligan et al. (2004) use an instrumental variable strategy based on the changes in compulsory schooling laws and find a strong effect of education on voting. They also find that educated adults are more likely to discuss politics and associate with political groups.

The cited papers highlight the importance of education for the functioning of democracy, measured as political participation or stability of democratic governments. They do not deal with the quality of politicians, that is instead at central stage in the political selection literature.

2.2 Political selection

Political selection analyzes the determinants of the attraction and retention of high quality individuals into politics. The need of high quality politicians relates to the presence of incomplete contracts between voters and politicians. When candidates cannot fully commit into policies in advance, credibility can be achieved only when candidates are perceived as competent by voters (Osborne and Slivinsky, 1996; and Besley and Coate 1997). In this literature, competency is a characteristic that can be observed ex-ante like, for example, education or success in previous jobs. A starting point of all analyses is that entry into politics is endogenous and, therefore, the quality of candidates crucially depends on the structure of the political competition and the relative payoffs for holding a political office.

Quality of the candidate-citizens. – Besley and Coate (1997) present a model of political selection characterized by endogenous entry from a pool of candidate-citizens. In their model, each citizen can, in theory, decide to run for an elective office. Agents are heterogeneous in terms of (observable) abilities and there are increasing returns to quality in

the labor market. Entry decision is based on the comparison between the expected benefits to hold an office (if elected) and the opportunity cost given by a job in the private sector; the individual probability to be elected rises with abilities. As a result, politicians tend to be extratcted by the upper tail distribution of abilities and a larger pool of high quality individuals mechanically generates an increase in the probability to elect a skilled politician.

Economic conditions and rent extraction. – Opportunity costs play a fundamental role in the literature. Caselli and Morelli (2004) address the issue by analyzing the probability to elect "bad" (in terms of observable characteristics) politicians. When returns to quality are particularly high, good politicians may decide to stay out of politics thus leaving room for lower quality candidates.[5] This effect is amplified when the observable characteristic is particularly noisy in signalling the true quality of an individual. The possibility for rent extraction is also a main driver for the political selection. Brollo et al. (2012) study the role of transfers to local authorities from both a theoretical and an empirical point of view. They show that larger transfers increase the probability for mismanagement and rent extraction. This unduly increases the returns for being in office, relatively more for lower ability individuals due to their lower market wage. Eventually, the larger the transfers, the more sustained the entry for low-quality candidates, with a deterioration of the overall quality of the political competition.

Contestability. – Another crucial determinant is the degree of contestability in local elections. In this set of models, the electoral outcome is determined by swing voters that is by citizens that care more of the quality of politicians than of ideological platforms. Galasso and Nannicini (2011) analyze the allocation of candidates across districts by two competing parties and find that a better political selection is associated with the contestability of the district, measured by the importance of swing voters in determining the electoral result. This is confirmed in a subsequent paper (Galasso and Nannicini, 2012), in which they compare majoritarian and proportional elections and show that the majoritarian system increases the overall quality of politicians only for the most contestable districts.

[5] Dal Bò et al. (2002) have extended the Caselli-Morelli analysis to give a role to pressure groups.

All-in-all this paper operates a conjunction between these two streams of literature. It shows a new linkage between education and the functioning of democracies and it adds a new determinant to the political selection issue.

3. How do (did) Italians elect their mayors?

In March 1993 the Italian Parliament approved an electoral reform that introduced the direct election of mayors (Law No. 81, 25 March 1993, *Elezione diretta del sindaco e del presidente della provincia*). New rules radically changed the existing political arena for the election of mayors and represented a drastic policy shift.

Pre-reform politics was characterized by a parliamentary system. The municipal council (*consiglio comunale*) was the only institutional body that was directly elected with a proportional system. Candidates' lists were entirely chosen by local sections of national parties as primary elections were completely unknown in the political system of that period. In many cases, majorities in the *consiglio comunale* were created after elections. The choice of the mayor was based on an ex-post bargaining among the local political party leaders; as a standard practice, mayors were not the local party leaders and were chosen among the available councilors. The system was characterized by a strong intermediation by political parties. It was nearly impossible to identify before elections a possible mayor, since post-election bargaining was particularly delicate and no party could seriously commit to choose a possible candidate. This was an intended outcome: the system was aimed at ensuring a close party control over local governmental activities and – at the same time – at preventing the mayor from consolidating any degree of autonomous leadership.[6] The Italian politics was, at that time, characterized by a low level of *ex-ante identifiability*. By analyzing national government formation, Strom (1990) points out that it was nearly impossible, before elections, to predict the winning candidate for the Prime Minister office; Strom also notes that "Many Italian governments have lacked of precommitted legislative majorities". Compared with other western democracies, the *ex-ante identifiability* was the lowest in his

[6] This system was actually mirroring the central government one, established with the 1948 Consitution with the stated aim to prevent the creation of strong executive power as a clear reaction to the 1922-1943 fascist regime.

sample.[7] As a result, the intermediation by political parties was overwhelming and voting was predominantly ideological as little room was left for the visibility of local politicians.[8]

The 1993 reform was set in a period of great transformation of the Italian politics. From 1990 on, past political paralysis, massive government debt, extensive corruption, and organized crime's considerable influence created a great demand for political, economic, and ethical changes. The reform was introduced to reinvigorate local governments and to simplify the administration of cities and provinces. The target was to meet a larger demand for accountability and limit the overwhelming influence of political parties.

Accountability was reached by establishing the direct election of mayors, with run-off for the two most voted candidates when no one reached 50% of ballots in cities larger than 15,000. Mayors were allowed to run for a second term and this represented a test for their political actions. Direct elections were associated with new powers, such as the possibility to recruit aldermen from outside the council or the obligation to proceed to new elections if mayors decide to resign. A limit in the national parties' influence was set for municipalities smaller than 5,000, in which candidates were supposed to create their own lists (so called *Liste Civiche*) without any formal affiliation to political parties. In larger than 15,000 municipalities, a substantial premium was granted to the winning candidate in the city council.

Compared with the previous system, the reform introduced a new (and more stable) system of representation. Voters had more influence and more incentives to carefully select their Mayors. This result was summarized by Baldini and Legnante (1998) according to which the new system secured a rather visible and independent position for the new mayors: "by being directly elected above and beyond the party wrangling in the communal council

[7] On similar dynamics, at local level, see also Agosta (1999), Parisi (1984), Pasquino (2001), and Vandelli (1997).

[8] It is not straightforward to understand which kind of political selection was at work in the pre-reform system. The paper by Mattozzi and Merlo (2012), however, is able to give some guidance on how party select their members and the possible candidates for elections. They find that parties may have weak incentives to select individuals with very high abilities since a relatively more homogeneous group (without superstars) is likely to maximize the collective effort of the party. In other words, selection by parties leads to a set of candidates whose characteristics are much closer to the median voter's characteristics.

and bureaucratic fetters, the mayor gains individual legitimation and visibility which are valuable instruments not so much for governing, but particularly for re-election."

A first, non causal, assessment for the effects of the reform is reported on figure 1. On the y-axis I report a measure of the quality of elected mayors in the years before and after the reform. The measure is computed as the difference between the years of schooling of the mayor in the year of election and the average schooling of all elected mayors in the same city between 1985 and 1997.[9] The figure seems to suggest that, after 1993, directly elected mayors were characterized by higher education, thus hinting that the 1993 reform actually increased the incentives to vote for higher quality candidates. The figure also shows a relevant improvement in the quality of mayors before the 1993 in the period 1990-92. This suggests that the increasing turmoil in the Italian politics due to the fall of the Berlin wall generated a rising demand for political changes. I will exploit this feature in a robustness check.

4. Empirical design

As explained in the Introduction, a typical way to identify a causal effect in the human capital externalities literature is to use some institutional factors that determines (exogenous) differences in the local human capital accumulation as an instrument for the current endowment of city-average schooling (Milligan et al., 2004; Acemoglu and Angrist, 1999). In this paper, I make use of an alternative strategy based on the increasing incentives to elect a high-quality mayor after the 1993 reform in a diff-in-diffs framework. If the effects of direct voting are heterogeneous according to the human capital endowments of the population and human capital is pre-determined with respect to the reform, I am able to estimate a causal parameter.

More formally the baseline model is as follows:

$$y_{ct} = \alpha_c + \delta_1 TIME_t + \delta_2 School_c * TIME_t + \xi_{ct} \qquad (1)$$

[9] This is done to take into account all the other confounding factors that may influence the political selection. See section 2.

where y_{ct} is a measure of quality of mayors elected in city *c* at time *t*. α_c is a city-level fixed effect, $TIME_t$ is a dummy variable equal to one for all years after 1993, and $School_c$ is a measure of educational attainment (human capital) at city level.

In order to correct the estimates for potential serial correlation within municipalities, I use one of the solutions proposed by Bertrand et al. (2004), that is to average the data before and after the reform and, therefore, to run a regression over two points in time. Pre-reform years are 1987-92; post-reform years are 1993-97.[10] Therefore I estimate equation (1) as follows:

$$\bar{y}_{ct} = \alpha_c + \beta_1 POST_t + \beta_2 School_c * POST_t + \varepsilon_{ct} \qquad t = 0,1 \qquad (2)$$

where $\bar{y}_{ct}$ is the average of the dependent variable before and after the reform and $POST_t$ is a dummy equal to one for the period after the reform.[11]

When estimating equation (2) I cluster standard errors at city level to take into account correlation across time within the same city.[12] The estimation of equation (2) with city-level fixed effects is the analog of taking first differences at municipality level.

Period fixed effects are meant to capture common aggregate trends that each year hit all municipalities in the same way. City fixed effects control for characteristics of each municipality that do not vary across time and might be correlated with the likelihood to appoint a high-quality politician. The interaction between $School_c$ and $TIME_t$ aims at capturing the effect of human capital on the quality of elected mayors after the reform. It can be considered as a causal parameter if there are not other omitted variables that correlate with both the schooling level of the population and, at the same time, the likelihood to elect a high-quality mayor after 1993.

[10] Five years were the maximum length of a local legislature.

[11] It should be noted that it is not possible to simply restrict the sample to the years just before and after the reform. This is due to the fact that elections do not take place every year and just a few observations are available for those two years.

[12] The use of alternative cluster or generic corrections for heteroskedasticity does not change the results.

There can be several sources of this kind of omitted variable biases. I will discuss them in details in the robustness check section (6.2).

5. Data

Data comes from two main sources: the Registry of local Administrators (RLA) and the 1991 and 1971 population censuses.

RLA is managed by the Central Directorate of Electoral Services of the Italian Ministry of the Interior. The registry is a picture of all elected officers in all Italian municipalities, provinces, and regions at the December 31st of each year starting from 1985. For municipalities, the definition of officer includes Mayor, vice-Mayor, Alderman, President, vice-President of the city council, and all councilors. For cities in which a Mayor is not available,[13] the database also indicates the Commissioner appointed by the Ministry of the Interior before new elections. The database is quite rich. It indicates political party affiliation, date and city of birth, educational attainment, and previous occupation. It also includes the year of election and the reason why she left her office.

RLA provides for the dependent variable of equation (2).

Defining the *quality* of a mayor is not an easy task. According to the political economy literature, *quality* is a characteristic that can be observed (sometimes imprecisely) ex-ante by voters. For this reason I use four measures that can be easily observed before elections and are usually correlated with the candidates' capabilities.

The first and most important is years of schooling, that is computed by attributing, as usual in the estimates of Mincerian equations for Italy, to each individual the lowest number of years necessary to obtain her higher degree. That is: zero years for no formal education (uncompleted elementary school); 5 years for elementary school; 8 years for lower secondary school; 13 years for upper secondary school; 18 years for university degree. The second indicator, strictly related to the previous one, is the share of elected mayors with at

[13] This can occur for several reasons: resignation, death, forced dissolution of the city council due to suspected organized crime infiltration.

least a secondary school degree. The third is experience in the labor market, that is calculated, as usual in the estimates of Micerian equations in Italy, as the difference between the age in the year of election and the number of years of schooling (de Blasio and Di Addario, 2005) plus 6.[14] The fourth is the skill intensity of the mayor's previous occupation. Skilled workers are: professors, journalists, physicians, pharmacists, judges, lawyers, engineers, chemists, biologists, architects, mathematicians, economists, statisticians, business consultants, and managers.[15]

The other main data sources are the Italian censuses of 1991 and 1971. They provide for the explanatory variable of interest, that is the educational level of resident population (aged at least 6) at city level. 1991 data is used in the baseline regression, while 1971 data is used as an instrument in a robustness check. Also in this case, I calculate the average schooling of the population by using the lowest number of years necessary to obtain the higher degree.

Table 1 reports the descriptive statistics of all variables used in the empirical analysis.

Before 1993 the average years of schooling of mayors was 13.3, that barely corresponds to the achievement of a high school diploma and is more than twice of the average years of schooling of population in 1991 (6.5). After the 1993 reform, the average years of schooling of mayors rose to 14.1. A similar increase is detectable for the other measures of quality of mayors: the share of mayors with at least a high school diploma grew from 78% to 84%; the portion of mayors with a high skilled job rose by 3 percentage points. A contrasting effect appears, instead, in the measure of experience in the labor market.

6. Results

In this section, I show the results by assessing whether human capital externalities have an effect the quality of political selection at city level.

[14] Italian schooling system starts at 6.

[15] For a similar definition see Brollo et al. (2012).

The core of my identification strategy is the diff-in-diffs estimation presented in section 4. In section 6.1 I show the baseline results on the average schooling of elected mayors. In section 6.2 I make a number of robustness checks that take into account all possible confounding factors that are likely to influence the quality of elected politicians (quality of the pool of citizen-candidates, population size, local economic conditions, and political contestability). Further robustness is made by considering alternative time spans and by making placebo experiments. Finally, I show that the main result holds when using alternative measures of quality of politicians.

6.1 Baseline results

By estimating equation (2) without the interaction term and using the log of the average schooling of elected mayors at city level before and after the 1993-reform as a measure of quality, I obtain a point estimate of β_1 of 0.064 (with a s.e. of 0.004, not reported, available upon request). This means that, coeteris paribus, the reform increased the education of elected mayors by 6.4%.

Estimates of equation (2) with the interaction terms are displayed in table 2. Column 1 reports the estimate for the full sample that includes all Italian cities. The coefficient of interest (β_2) is positive and significant with a point estimate of 0.078. This means that, for a municipality at the 75th percentile of the human capital distribution, the reform raised the quality of elected mayor by 7 per cent while the effect for a municipality at the 25th percentile was 5.8 per cent. This implies that the interquartile range in the human capital distribution accounts by almost one fifth ((7-5.8)/6.4 = 0.18) of the total average increase after the reform. This is a sizable effect if we consider the radical transformations in Italian politics in the period under consideration.

In columns (2) and (3) I take into account some issues related to city size. The presence in the same sample of both very small and very large cities is likely to be a source of confounding factors for both institutional and political reasons.

From an institutional point of view (see par. 3), the 1993 reform established different electoral rules whether the city is smaller or larger than 15,000. For smaller municipalities, all elections are with a first-past-the post system, and the winner is the one who gains more

votes. For cities larger than 15,000, electoral rules contemplate a second round if none of the candidates reached the 50% of votes in the first round. This is likely to change the setting of the political competition in those cities.

From a political point of view, city size may matter a lot in the perception of the local results. Mayoral elections are likely to have national consequences if the city at stake is particularly large. This means that political campaigns in larger municipalities might have different characteristics as country-wide leaders of national parties could have strong incentives to intervene.

In order to address these issues I run equation (2) on a subset of smaller cities. I chose two population thresholds. For the institutional issue, I focus on the single-round municipalities, that is those smaller than 15,000. For the political issue, I chose a very restrictive threshold of 5,000: this is due to the fact that in municipalities smaller than that size, candidates cannot be affiliated to national parties but should run with their own party list (the so-called *liste civiche*), thus reducing the potential influence of national politics.

Estimates are provided in columns (2) and (3) for, respectively, the 15,000 and 5,000 inhabitants threshold. The baseline result is confirmed: estimated value of β_2 is positive and significant and even slightly larger (0.087) than the baseline one.[16]

6.2 Robustness

6.2.1 Besley and Coate argument and IV estimates

Diff-in-diffs specifications are able to capture a causal effect only in absence of other confounding factors that are correlated with both the educational level of the population and the probability to elect a high-quality mayor after the reform.

The first challenge relates to the quality of the local pool of candidate citizens. In Besley and Coate (1997), candidates are generally a subset of voters; this implies that the more educated is the pool of citizens, the higher the probability to have a high quality candidate. The Besley-Coate argument is a great challenge for the causal interpretation of

[16] The implied effects for both thresholds are quite similar to the baseline one as the interquantile ranges for the restricted samples are slightly smaller than the full sample.

the estimates; in their model all citizens care in the same way of the quality of the politicians and the positive correlation between the quality of mayors and human capital endowment is mechanical. The diff-in-diffs specification is particularly powerful under this respect. If the schooling of population at city level is time-invariant, Besley-Coate argument is fully absorbed by city-level fixed effects.

However, in the data, the human capital endowment does not need to be constant over time as Italian law does not forbid within country migrations. Relocation decisions by skilled citizens into more educated cities when the reform took place may create an upward bias in the estimates.

A possible way to cope with this problem is to use a time-varying measure of human capital at city level. This solution cannot be adopted in this context for two reasons. The first is that a time-varying measure of schooling would be endogenous, thus generating an inconsistency in the estimates of the parameter. The second is that a complete survey on the educational level of population is conducted on censuses only, thus implying that changes in the human capital endowment at city level is unobservable to the econometrician.[17]

In order to cope with this problem, I instrument the 1991 average schooling of the population with its 1971 level. This is a good instrument as long as the 1993 reform could not be foreseen in 1971, that is a not particularly heroic assumption.

Results are displayed in table 3. First stage statistics are quite satisfactory: the F of excluded instruments is very high and well above the rule-of-thumb value of 10. The coefficient of the first stage is also positive and highly significant, thus indicating a good correlation between current and past values of educational levels at city level. Second stage results are also quite reassuring. The coefficient is positive, significant, and quite similar in magnitude with respect to the OLS values, with the only exception of the estimate for cities smaller than 5000, that is slightly larger.[18]

[17] The following population census was conducted in 2001.

[18] This is not surprising since migration impact more on smaller municipalities.

6.2.1 Other omitted variables: contestability, economy, and city size

Another source of bias in equation (2) is a change in the conditions of political competition. The reform was introduced in a period of increasing contestability of the Italian political system, due, mainly, to the access to the political arena of both left- and right-winged new political parties. As explained in section 2, higher contestability generally entails the election of higher quality politicians. If the increase in contestability were equally spread across all municipalities, its effect would be simply captured by β_1. If, instead, it were more pronounced in more educated cities, estimates would be upward biased.

In order to tackle this issue, I insert a measure of political contestability taken from general elections results (*Margin of victory*).[19] I construct a variable that is the difference in the share of votes between of the winning and the second party at city level in the general elections of 1987 and 1992 (for the years before) and 1994 and 1996 (for the years after). For all the available years, I took the votes in the proportional elections.[20] Preliminary evidence on the strong increase in the contestability is shown in table 1. For the pre-reform elections (1987 and 1992) the average margin of victory was 20 percentage points. In the subsequent two elections (1994 and 1996), it shrank to 11.

Another relevant confounding factor is the heterogeneous change of skill premia in Italian cities after the reform. The early 90s were characterized by a strong diffusion of skill biased technologies that were likely to change the returns to education and, hence, the incentives for more educated individuals to participate to electoral competitions. If skill-biased technologies were more likely to be adopted in cities with higher human capital (Acemoglu, 1998; Accetturo et al., 2013), I would observe an attenuation bias in the estimate of β_2.

[19] Data are taken from the Istituto Carlo Cattaneo (ICC) database on national Italian elections from 1861 to 2008. ICC is a private research organization aimed at promoting studies on political issues in Italy. ICC collects a huge database on results of political elections at municipal level.

[20] This is a relevant issue for the 1994 and 1996 general elections. For those years, a mixed electoral system was at work. Part of the members of Parliament was elected in majoritarian consituencies, while the remaining part was elected with a proportional system. All voters were given two voting papers, one for the majoritarian vote and the other for the proportional vote. In the calculation of this measure, I only took the votes for the proportional seats. All elections before 1994, instead, were done with a purely proportional system.

In order to capture this possible confounding factor, I use the 1991 and 1996 economic censuses to control for changes in the labor demand for more educated workers (*log empl_ser*). In particular, I insert a control variable equal to the log of employment in high-skilled sectors for the two census years. High skilled industries are selected by using the EU-KLEMS database and are defined as the industries that, over the period 1991-96, employed a share of individuals with a university degree higher than the Italian average. These industries are: Education (Nace rev. 1: M), Health (N), Financial Intermediation (J), Renting and Business activities (K), and other public administration (L).

I finally add population size as regressor (*log POP*). This is aimed at controlling for the fact that a change in the institutional settings may have different results according to the size of the constituency. As the theory of electoral turnout shows (Taylor and Yildirim, 2010), the probability to be pivotal in a winner-takes-all elections shrinks with population size, with relevant effects on voting behavior. Population size, instead, has a more marginal effects in proportional elections (Herrera et al., 2012) as those before the reform.

Results are displayed in Table 4 and confirm the baseline estimates of Table 2. The measure of contestability is never statistically significant, thus suggesting that changes in contestability equally affected all municipalities. Also the controls for high-skilled sectors and the log of population are not significant.

Although these controls can actually capture some of the changes in the nature of political competition before and after the reform they may not fully control for all remaining confounding factors. I assess the relative importance of unobservable omitted variables by analyzing the possible variation in the coefficient of interest change with the inclusion of explanatory variables. If additional controls substantially attenuate the estimates, it is possible that inclusion of more controls would drive the estimated effect to zero. Conversely, if the inclusion of controls does not change in a relevant way the point estimate, I can more confidently claim the causal interpretation of the parameter.

Following Altonji et al. (2005) and its extension to the continuous case made by Bellows and Miguel (2009), I measure of the relative strength that omitted variables should have relative to the observed controls to completely wash away my result. This is calculated as the ratio between the coefficient of interest with controls (for example, 0.064 for the full

sample case) and the difference between the coefficient without controls (0.079, this is the coefficient on the same estimation sample, as reported at the bottom of the table) and the coefficient with controls (0.064). Results for these calculations are reported in the last row of the table. If the set of observed controls is representative of all possible controls, then a large ratio suggests that it is implausible that omitted variable bias explains away the entire effect. In the full sample case, unobserved omitted variables should be 4.26 times stronger than observed controls to drive the coefficient to zero, which seems highly unlikely.[21]

6.2.3 Alternative time spans and placebo experiments

I also further explore the robustness of the results by changing the time span of analysis. As figure 1 shows, a substantial process of quality upgrading in the elected mayors was in place in the period 1990-92, probably due to increasing competitive pressure in the political arena. The observed increase in schooling for elected mayors in more educated cities may, therefore, reflect a smooth process of quality upgrading that is totally unrelated with human capital externalities issues. In order to cope with this problem, in table 5, I restrict the analysis to the period 1990-96; I use the 1990-92 years (in which a rise in the quality was already at place) as pre-reform and 1993-96 as post-reform. Estimated coefficients are now slightly larger although confidence intervals largely overlap with the baseline result. This implies that the rise in the quality of elected mayors is not related to a smooth transition from a less to a more competitive political system, but it is entirely due to a change in the incentives to elect better politicians after the reform.

Another way to rule out this issue is to perform a placebo experiment. In table 6, I choose an arbitrary faked reform year and analyze whether I observe any change in the quality of elected politicians in that year. The faked reform year is 1990. From a political point of view, this is a crucial year since it follows the fall of the Berlin wall and, therefore, the end of the Cold-war era Italian political system. In this case, pre-reform period is 1987-89 and post-reform is 1990-92. Results are displayed in Table 6: coefficients are now smaller and quite far from standard significance levels. Results in Table 6 are particularly powerful

[21] Altonji et al. (2005) consider 3.55 a quite safe ratio for their estimates.

from another perspective: they show that before the reform, high human capital and low human capital cities displayed similar patterns in the evolution of political selection. In other words, it shows that the common trend pre-treatment assumption is respected (Angrist and Pischke, 2009) and, therefore, the causal interpretation of a diff-in-diffs estimator is warranted.

Another placebo experiment is done by analyzing the effects of the reform on the average schooling of councilors. As explained in section 3, the 1993 reform affected the way mayors were elected but they dealt only marginally with city councils. In particular, for cities greater than 15,000, a substantial majority premium was granted to the winning major, whereas nothing changed before and after the reform for cities smaller that that threshold. Results are displayed in table 7. The coefficient of interest in columns (2) and (3), i.e. for councils unaffected by the reform, is far from standard significance levels, thus showing that political selection did not significantly change in more educated cities after the reform. The coefficient of column (1), that includes municipalities with majority premia, is instead negative and significant, thus indicating a negative political selection in more educated cities. This is reasonable in a context in which city councils lost relevance with respect to mayors and the hiring of educated councilors is costly for the party (see Galasso and Nannicini, 2011, for a similar argument). All-in-all it should be noted that the period dummy is always positive and significant, thus indicating a general skill upgrading in the quality of politicians after 1993, probably due to the stronger demand for political change after the early-90s scandals.

6.2.4 Alternative measures

All regressions run so far rely on a single proxy for the quality of politicians (education). In table 8, I report the baseline estimates for other measures that are commonly used in the empirical literature on political selection: average years of experience, share of mayors with a high-skilled job, share of mayors with at least a secondary education. Results show that even by using alternative definition, the baseline result is confirmed.

7. Concluding remarks

This paper analyzes a new area over which education may exert its effects: political selection. Most of the existing literature has studied the social returns to education in the labor market, criminal activity, and political participation. However, the fact that education increases the cognitive skills, lowers the costs to grasp information, and raise the capability to understand the issues at stake should have an effect also on the ability to select more able politicians.

This paper is aimed at analyzing whether cities with a larger endowment of human capital tend to elect mayors with ex-ante better characteristics. This is done by exploiting an electoral reform for the election of Italian mayors that generated stronger incentives to evaluate the mayor's characteristics in a diff-in-diffs framework. Results show that more educated citizens tend to elect higher quality politicians.

This result is robust to omitted variable and citizen sorting biases; it is confirmed by using alternative measures of quality. The evidence is consistent with a causal interpretation from higher education to better ability to select higher-quality politicians.

References

Accetturo A., A. Dalmazzo and G. de Blasio, 2013, "Skill Polarization in Local Labour Markets under Share-Altering Technical Change", forthcoming *Journal of Regional Science*.

Acemoglu, D., 1998. "Why Do New Technologies Complement Skills? Directed Technical Change And Wage Inequality," The Quarterly Journal of Economics, MIT Press, vol. 113(4), pages 1055-1089, November.

Acemoglu, D. and J. Angrist, 1999. "How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws," NBER Working Papers 7444, National Bureau of Economic Research, Inc.

Agosta, A., 1999, 'Sistema elettorale e governo locale: Gli effetti politici e istituzionali della riforma del 1993', in S. Operto (ed.) Votare in città. Riflessioni sulle elezioni amministrative, Milan, Angeli, pp. 31–58.

Almond, G. and S. Verba, 1989. "The Civic Culture: Political Attitudes and Democracy in Five Nations," Sage Publications.

Altonji J.G., Elder T., Taber C., 2005, "Selection on observed and unobserved variables: assessing the effectiveness of catholic schools" Journal of Political Economy, 113 (1) (2005), pp. 151–184.

Baldini Gianfranco and Guido Legnante (1998): «The 1997 Local Elections in Italy: From party Mayors to a Mayors' Party», in L.Bardi and M.Rhodes (eds): Italian Politics. Oxford: Westview Press.

Barro, R. (1999). Determinants of democracy. Journal of Political Economy, 107, 158–183.

Bellows J., Miguel, E., 2009, "War and local collective action in Sierra Leone" Journal of Public Economics, vol. 93, n. 11-12, pp. 1144-1157.

Bertrand, M.; Duflo, E.; Mullainathan, S. (2004) "How Much Should We Trust Differences-in-Differences Estimates?" The Quarterly Journal of Economics, 119 (1), 249-275

Besley, Timothy and Coate, Stephen, 1997. "An Economic Model of Representative Democracy," The Quarterly Journal of Economics, MIT Press, vol. 112(1), pages 85-114, February.

Bourguignon, F., and Verdier, T. (2000). Oligarchy, democracy, inequality, and growth. Journal of Development Economics, 62, 285–313.

Brollo, F., T. Nannicini, R. Perotti and G. Tabellini. 2012. The political resource course. Forthcoming, American Economic Review.

Caselli, Francesco and Massimo Morelli. 2004. "Bad Politicians." Journal of Public Economics. 88: 3–4, pp. 759–82.

Dal Bo, E., Dal Bo, P., Di Tella, R., 2006. Plata o Plomo? Bribe and Punishment in a Theory of Political Influence. American Political Science Review 100(1), 41–53.

De Blasio, G. and Di Addario, S. (2005), Do Workers Benefit from Industrial Agglomeration?. Journal of Regional Science, 45: 797–827.

Friedman M., 1962. “Capitalism and Freedom”.

Galasso V. and Nannicini T., 2011, "Competing on Good Politicians", American Political Science Review, 105, 79-99.

Galasso V. and Nannicini T., 2012, "Political Selection under Alternative Electoral Rules", mimeo.

Glaeser, E., LaPorta, R., Lopez-de-Silanes, F., and Shleifer, A. (2004). Do institutions cause growth? Journal of Economic Growth, 9, 271–303.

Glaeser, Edward L. and Gottlieb, Joshua D., 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," Journal of Economic Literature, American Economic Association, vol. 47(4), pages 983-1028, December.

Glaeser, Edward and Ponzetto, Giacomo and Shleifer, Andrei, 2007. "Why does democracy need education?," Journal of Economic Growth, Springer, vol. 12(2), pages 77-99, June.

Herrera, H. and Morelli, M. and T. Palfrey, 2012, “Turnout and Power Sharing”, mimeo.

Lipset, S. M. (1960). Political man: The social basis of modern politics. New York: Doubleday.

Lorchner, L., 1999. "Education, Work, and Crime: Theory and Evidence," Rochester School for Economic Research w.p. n. 465.

Lorchner, L. and Moretti, Enrico, 2002. "The effects of Education on Crime: Evidence from prison inmates, arrests, and self-reports," mimeo.

Mattozzi Andrea and Merlo Antonio, 2011. "Mediocracy, Fourth Version," PIER Working Paper Archive 13-010, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania, revised 08 Feb 2013.

Milligan, Kevin and Moretti, Enrico and Oreopoulos, Philip, 2004. "Does education improve citizenship? Evidence from the United States and the United Kingdom," Journal of Public Economics, Elsevier, vol. 88(9-10), pages 1667-1695, August.

Moretti, Enrico, 2004. "Human capital externalities in cities," Handbook of Regional and Urban Economics, in: J. V. Henderson and J. F. Thisse (ed.), Handbook of Regional and Urban Economics, edition 1, volume 4, chapter 51, pages 2243-2291 Elsevier.

Osborne, Martin J. and Al Slivinski. 1996. “A Model of Political Competition with Citizen Candidates.” Quarterly Journal of Economics. 111:1, pp. 65–96.

Papaioannou, Elias and Siourounis, Gregorios, 2008. "Democratisation and Growth," Economic Journal, Royal Economic Society, vol. 118(532), pages 1520-1551, October.

Parisi L. M., 1984, Luoghi e misure della politica. Bologna, Il Mulino.

Pasquino G., 2001, ,‘Le elezioni comunali nella transizione italiana: La costruzione della leadership’, Teoria Politica 17(2): 91–106;

Strom Kaare, 1990, Minority Government and Majority Rule, Cambridge University Press.

Taylor, C.R. and Yildirim, H., 2010, A unified analysis of rational voting with private values and group-specific costs, Games and Economic Behavior 70 (2), 457-471

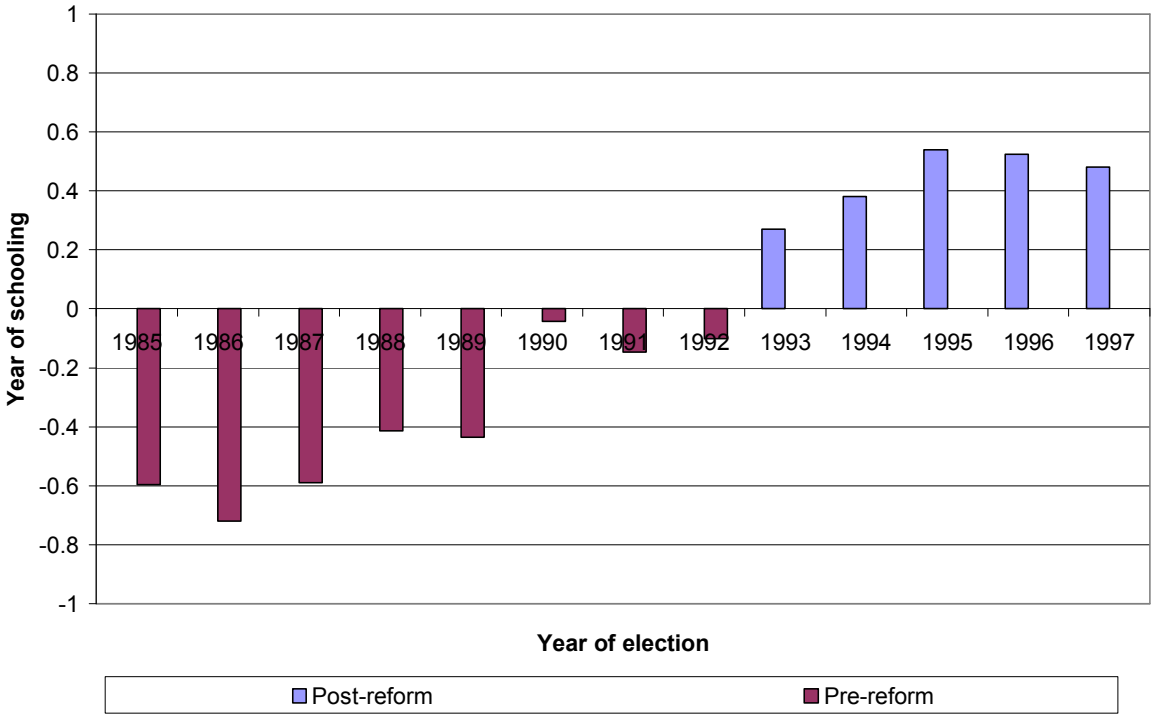
Vandelli L., 1997, Sindaci e miti. Sisifo, Tantalo e Damocle nell’amministrazione locale. Bologna, Il Mulino.

Verhoogen, Eric, 2008. "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," The Quarterly Journal of Economics, MIT Press, vol. 123(2), pages 489-530, 05.

Veronese, Barbara. 2005. “Representation, Policy Making and Accountability: Learning From Changes in Democratic Institutions.” Unpublished Ph.D. thesis, chapter 4, London School of Economics.

Figure 1

QUALITY OF MAYORS AND ELECTORAL REFORM



Source: Author's calculations on the Registry of local Administrators.

Table 1

DESCRIPTIVE STATISTICS

	No. Obs.	Mean	Standard Deviation
Years of Schooling – before 1993	7,332	13.330	3.719
Years of Schooling – after 1993	7,332	14.123	3.669
Years of Schooling – population (1991)	7,332	6.452	0.776
Years of Schooling – population (1971)	7,332	4.020	0.883
Population	7,332	7,137	43,817
Change of ruling party – before 1993	7,332	0.275	0.132
Margin of victory in national elections – before 1993	7,221	0.202	0.149
Margin of victory in national elections – after 1993	7,221	0.111	0.092
Years of experience – before 1993	7,332	26.103	10.225
Years of experience – after 1993	7,332	25.694	10.502
Share of high skilled jobs – before 1993	7,332	0.537	0.432
Share of high skilled jobs – after 1993	7,332	0.563	0.477
Share with secondary education – before 1993	7,332	0.776	0.377
Share with secondary education – after 1993	7,332	0.841	0.355
Log(empl)-services – 1991	7,332	3.388	1.777
Log(empl)-services – 1996	7,332	3.388	1.778
Years of Schooling (councilors) – before 1993	7,332	11.136	1.918
Years of Schooling (councilors) – after 1993	7,332	11.890	3.388

Source: Author's calculations on the Registry of local Administrators, 1971, 1991, and 1996 censuses, and ICC.

Table 2

BASELINE RESULTS

Dependent variable: log(schooling) – Mayor	Full sample	Pop<15,000	Pop<5,000
Log(schooling) – citizens (1991) * Post	0.078**	0.087**	0.087**
	(0.028)	(0.031)	(0.036)
Post	-0.082	-0.099*	-0.099*
	(0.053)	(0.058)	(0.067)
Constant	2.541***	2.529***	2.506***
	(0.001)	(0.002)	(0.002)
No. Obs.	14,664	13,482	10,626
R^2 – within	0.038	0.037	0.032
Number of clusters	7,332	6,741	5,313

Source: Author's calculations on the Registry of local Administrators and 1991 censuses.
FE estimates. Clustered standard errors in parenthesis at city level. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 3

IV ESTIMATES

Dependent variable: log(schooling) – Mayor	Full sample	Pop<15,000	Pop<5,000
Log(schooling) – citizens (1991) * Post	0.072**	0.087**	0.125**
	(0.039)	(0.044)	(0.050)
Post	-0.070	-0.098	-0.169*
	(0.072)	(0.080)	(0.093)
No. Obs.	14,664	13,482	10,626
R^2 – within	0.038	0.037	0.032
Number of clusters	7,332	6,741	5,313
F-first stage	9,144	7,904	9,144
Marginal effect (instrument)	0.412***	0.397***	0.385***
	(0.004)	(0.004)	(0.005)

Source: Author's calculations on the Registry of local Administrators and 1971 and 1991 censuses.
IV estimates. Instrument: log of the citizens' average schooling in 1971.
Clustered standard errors in parenthesis at city level. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 4

CONTROL FOR CONTESTABILITY, POPULATION, AND ECONOMIC CONDITIONS

Dependent variable: log(schooling) – Mayor	Full sample	Pop<15,000	Pop<5,000
Log(schooling) – citizens (1991) * Post (A)	0.064**	0.071**	0.080**
	(0.031)	(0.033)	(0.038)
Margin of victory	0.031	0.027	0.014
	(0.026)	(0.027)	(0.030)
Log(pop)	0.001	0.002	0.000
	(0.003)	(0.004)	(0.005)
Log(empl_ser)	0.007	0.009	0.006
	(0.007)	(0.007)	(0.007)
Post	-0.060	-0.081	-0.086
	(0.057)	(0.066)	(0.084)
Constant	2.513***	2.498***	2.489***
	(0.026)	(0.024)	(0.021)
No. Obs.	14,505	13,328	10,479
R^2 – within	0.038	0.036	0.032
Number of clusters	7,293	6,704	5,279
Result of baseline estimates on the same estimation sample			
Log(schooling) – citizens (1991) * Post (B)	0.079**	0.089**	0.088**
	(0.029)	(0.031)	(0.037)
(A)/[(B)-(A)]	4.66	3.94	10.00

Source: Author's calculations on the Registry of local Administrators, 1991 and 1996 censuses, and ICC.
FE estimates. Clustered standard errors in parenthesis at city level. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 5

TIME SPAN: 1990-96

Dependent variable: log(schooling) – Mayor	Full sample	Pop<15,000	Pop<5,000
Log(schooling) – citizens (1991) * Post	0.095**	0.102**	0.100**
	(0.032)	(0.035)	(0.041)
Post	-0.123**	-0.138**	-0.133*
	(0.060)	(0.065)	(0.075)
Constant	2.546***	2.534***	2.509***
	(0.001)	(0.002)	(0.002)
No. Obs.	14,664	13,482	10,626
R^2 – within	0.026	0.024	0.021
Number of clusters	7,332	6,741	5,313

Source: Author's calculations on the Registry of local Administrators and 1991 censuses.
FE estimates. Clustered standard errors in parenthesis at city level. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 6

PLACEBO: 1987-92

Dependent variable: log(schooling) – Mayor	Full sample	Pop<15,000	Pop<5,000
Log(schooling) – citizens (1991) * Post	-0.030	0.021	0.071
	(0.061)	(0.078)	(0.110)
Post	0.094	0.006	-0.071
	(0.114)	(0.143)	(0.199)
Constant	2.581***	2.558***	2.515***
	(0.004)	(0.004)	(0.007)
No. Obs.	14,664	13,482	10,626
R^2 – within	0.013	0.018	0.027
Number of clusters	7,332	6,741	5,313

Source: Author's calculations on the Registry of local Administrators and 1991 censuses.
FE estimates. Clustered standard errors in parenthesis at city level. Placebo discontinuity is set in 1990. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 7

PLACEBO: EFFECTS ON COUCILLORS

Dependent variable: log(schooling) – Councillors	Full sample	Pop<15,000	Pop<5,000
Log(schooling) – citizens (1991) * Post	-0.050***	-0.018	-0.004
	(0.012)	(0.013)	(0.015)
Post	0.161***	0.108***	0.094**
	(0.022)	(0.024)	(0.028)
Constant	2.395***	2.376***	2.334***
	(0.000)	(0.001)	(0.001)
No. Obs.	14,628	13,446	10,591
R^2 – within	0.233	0.254	0.298
Number of clusters	7,332	6,741	5,313

Source: Author's calculations on the Registry of local Administrators and 1991 censuses.
FE estimates. Clustered standard errors in parenthesis at city level. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 8

ALTERNATIVE MEASURES

	Full sample	Pop<15,000	Pop<5,000
	Years of experience		
Log(schooling) – citizens (1991) * Post	0.098*	0.096*	0.098
	(0.050)	(0.055)	(0.062)
Post	-0.212**	-0.208**	-0.201*
	(0.095)	(0.102)	(0.115)
Constant	3.181***	3.182***	3.192***
	(0.003)	(0.003)	(0.003)
No. Obs.	14,664	13,482	10,626
R^2 – within	0.003	0.003	0.003
Number of clusters	7,332	6,741	5,313
	Share of high-skilled jobs		
Log(schooling) – citizens (1991) * Post	0.158**	0.144**	0.118**
	(0.048)	(0.051)	(0.059)
Post	-0.267**	-0.245**	-0.202*
	(0.089)	(0.095)	(0.108)
Constant	0.537***	0.525***	0.503***
	(0.003)	(0.003)	(0.003)
No. Obs.	14,664	13,482	10,626
R^2 – within	0.003	0.003	0.003
Number of clusters	7,332	6,741	5,313
	Share of secondary education		
Log(schooling) – citizens (1991) * Post	0.093**	0.110**	0.112**
	(0.032)	(0.036)	(0.042)
Post	-0.107*	-0.138**	-0.143*
	(0.060)	(0.067)	(0.076)
Constant	0.776***	0.776***	0.746***
	(0.002)	(0.002)	(0.002)
No. Obs.	14,664	13,482	10,626
R^2 – within	0.030	0.030	0.027
Number of clusters	7,332	6,741	5,313

Source: Author's calculations on the Registry of local Administrators and 1991 censuses.
FE estimates. Clustered standard errors in parenthesis at city level. * significant at 10%, ** significant at 5%, *** significant at 1%.