

Manager or professional politician? Local fiscal autonomy and the skills of elected officials*

Massimo Bordignon[†]

Catholic University, Milan & CESifo

Matteo Gamalerio

Institut d’Economia de Barcelona (IEB), University of Barcelona

Gilberto Turati

Catholic University, Rome

12th February 2019

Abstract

We provide a theoretical and empirical assessment of why local fiscal autonomy can affect the skills of elected officials in sub-national governments. We first develop a model of politics with different types of politicians and show that – following a tax decentralization reform increasing local fiscal autonomy – politicians with high administrative skills are elected in rich jurisdictions while politicians with high political skills are elected in poor ones. As a result, voter welfare increases only, or mainly, in rich jurisdictions. We then look for empirical support to these predictions by exploiting the decentralization reforms affecting Italian municipalities in the '90s. These reforms introduced both the direct election of the mayor and new autonomous tax tools for municipalities characterized by large differences in their tax bases. Our estimates – robust to several alternative stories – emphasize a differential change in elected officials at the municipal level between rich and poor jurisdictions. These findings provide a new explanation for the observed poor performance of local governments largely financed by grants.

Keywords: decentralization, skills of politicians, fiscal autonomy.

JEL codes: D72, D78

*We wish to thank Fernanda Brolo, Peter Buisseret, Gianmarco Daniele, Philippe de Donder, Mirko Draca, Alessandro Fontana, Vincenzo Galasso, Anke Kessler, Ben Lockwood, Marko Koethenburger, Andrea Mattozzi, Giovanna Messina, Tommaso Nannicini, Giacomo Ponzetto, Francesco Porcelli, Panu Poutvaara, Marzia Romanelli, Paolo Sestito, Francesco Sobbrino, Guido Tabellini, Davide Ticchi, Dave Treisman, Andrea Vindigni, Jingjing Zhang for useful comments on previous drafts of this paper. Usual disclaimers apply. Declarations of interest: none.

[†]Corresponding author: Department of Economics and Public Finance, Catholic University, Largo A. Gemelli 1, I-20123 Milano, Italy. Ph.: +39.02.7234.2694; E-mail: massimo.bordignon@unicatt.it.

1 Introduction

The case for fiscal decentralization is in general terms rather weak. The theoretical literature emphasizes the efficiency gains, in terms of better representation of preferences and better accountability and promotion of politicians, that derive from devolving the provision of public services and the implementation of public policies to local governments (e.g., Myerson, 2006; Weingast, 2009; Lockwood, 2015). The same ideas lie behind the support that fiscal decentralization receives in political circles, international organizations and development agencies (e.g., Treisman, 2007; Bardhan and Mookherjee, 2006)¹. However, the empirical evidence is less supportive, with contrasting observed effects of decentralization in terms of efficiency, growth, quality of services, corruption, financial stability and the like (e.g., Rodden, 2002, 2006). Understanding under which conditions fiscal decentralization fulfills its promises is then crucial.

On this matter, a paramount role is played by the mismatch between own revenues and expenditures at the local level. Lower levels of fiscal autonomy (hence, higher shares of central government transfers in the local government budgets) are typically associated with poorer local governments' performances (e.g., Ahmad and Brosio, 2008). For instance, in communities largely financed by grants, local governments have a higher propensity to spend out of transfers, potentially leading to inefficient levels of public expenditure (the so-called "fly-paper effect"; e.g., Hines and Thaler, 1995, and Dahlberg et al., 2008), larger deficits (Rodden, 2002), and more corruption (Fisman and Gatti, 2002). Similarly, a substantial empirical evidence shows that financial instability and soft budget constraints problems are more likely to occur when fiscal autonomy is low (e.g., Rodden, 2002; Rodden et al., 2003; Bordignon and Turati, 2009; Eyraud and Lusinyan, 2013). Moreover, the quality of education provision and health care services at the local level is poorer in localities with lower financial autonomy (e.g., Reinikka and Svensson, 2004; Galiani et al., 2008; Francese et al., 2014). In addition, a large use of fiscal grants may lead to allocative inefficiencies, given that local governments politically aligned with the central government tend to receive more grants per capita (e.g., Levitt and Snyder, 1995; Larcinese, V., L. Rizzo, and C. Testa, 2006; Brollo and Nannicini, 2012).

However, why a lower degree of fiscal autonomy is associated with poorer performances of local government has never been made entirely clear in the literature. For instance, it is intuitive that local governments largely financed by transfers may not completely internalize the cost of spending, presumably because citizens have less incentives in controlling how much money is spent if a large part of it does not come from their own pockets². But it is still unclear why citizens should tolerate higher levels of corruption, or a lower quality of services, just because the latter are largely financed with grants. Similarly, under decentralization, local politicians may be more easily captured by local interests, leading to higher corruption or lower quality of services, as already indicated in the *Federalist papers*³. Again, it is not obvious why politicians should be more frequently captured when the level of financial autonomy is low.

In this paper we discuss an additional argument, which links the level of fiscal autonomy to

¹Treisman (2007) estimates in several hundred million dollars the total sum that *each year* international organizations, banks, development agencies, single states etc. donate or lend to developing countries in order to support decentralization.

²These are all instances of a "common pool" phenomenon (e.g., Persson and Tabellini, 1994, 2000), or the "1/n law" as it is alternatively defined in the legislative bargaining literature (Weingast et al. 1981). Cai and Treisman (2005) provide somewhat different theoretical arguments that point in the same direction.

³See Bardhan and Mookherjee (2000) and Bordignon et al. (2008) for modern treatments and discussion.

the *skills* of elected politicians in sub-national governments. We first develop a new theoretical model capturing the idea that politicians differ with respect to their skills – managerial vs. political skills – and the usefulness of these skills for voters depend on the characteristics of a jurisdiction. In particular, in a decentralized setting where most of the resources comes from the center, the main task of a local politician is to make sure that these resources keep flowing to the local community. This requires political skills (e.g., strong party connections with the center, particular bargaining abilities, and extended political networks) that are different from the skills of a good administrator of local matters. Thus, given the choice, voters of communities with lower levels of fiscal autonomy would rationally find more useful a politician with strong political connections rather than a good administrator. On the contrary, in communities where most of the resources are locally generated, voters would prefer a good administrator rather than a politician with good connections. The choice of voters can also be reinforced by a self-selection mechanism: anticipating voters’ preferences, different types of candidates enter politics in communities with different levels of fiscal autonomy.

The choice of these different politicians implies different consequences in terms of welfare: our model shows that tax decentralization reforms might have opposite effects in rich and poor communities. It increases voters’ welfare in rich communities, as it increases the level of fiscal autonomy and consequently attracts more politicians with high administrative skills, who are more efficient where fiscal autonomy is high. On the contrary, tax decentralization may reduce welfare in poor communities. This is because in these jurisdictions the degree of fiscal autonomy only slightly changes following a devolution of fiscal powers. Therefore, politically skilled politicians maintain a higher probability of being elected, even if their skills become less useful to voters. Interestingly, we prove that this may occur even with a “compensated” tax reform that leaves the total revenues of local governments unchanged, so that only the degree of fiscal autonomy changes with the reform.

We then look for an empirical support to our argument, exploiting the Italian experience of decentralization involving municipalities in the ’90s. Following political and economic turmoil at the beginning of the ’90s, two main reforms tried to improve the performance of local governments. In 1993 the central government devolved more fiscal powers to municipalities and changed the municipal electoral system. The fiscal reform introduced a new tax source, the municipal property tax, that had very different effects across the country. In the richest municipalities, because of their larger tax base, the newly introduced property tax made these municipalities almost entirely financially independent from the center. In the poorer ones, the effect was minimal and municipalities kept receiving most of their resources from the centre in the form of grants. The electoral reform introduced the direct election of mayors, assigning them a paramount role in municipal politics. Hence, the personal characteristics of candidates became more important to voters, who could now directly choose their main representative with the same rules in all jurisdictions. In 1999, the introduction of a municipal surcharge on the personal income tax of residents further enhanced this asymmetry between rich and poor cities (e.g., Bordignon, Grembi and Piazza, 2017). According to our theoretical model, following the reform package, we expect an abrupt change in the characteristics of politicians in the richer municipalities and a smaller or no effect in the poorer ones, followed by an improvement of welfare in richer communities only.

To test these predictions, we collect an extensive data-set on the personal characteristics of mayors of the main Italian cities both before and after the 1993 and the 1999 reforms, as well as on other economic and political features of the municipalities. Departing from the political economy literature interested in the “quality” of politicians, which considers either their level

of education or their income (e.g., Gagliarducci and Nannicini, 2013; Galasso and Nannicini, 2011), we use information on both mayor’s job before entering in politics and her/his former experience in politics to construct several proxy measures for different “types” of politicians, distinguishing between mayors with political skills and mayors with administrative/managerial skills. We also consider two indicators (the ratio between *collected* and *assessed* revenues and the percentage of “separate” waste collection) as proxies for voters’ welfare, which we expect to improve in rich communities. Besides being more reliable than budget data, none of the two indicators is directly related to the higher tax revenues of the richer cities.

Our evidence clearly shows a differential change in the characteristics of elected mayors between rich and poor municipalities following the 1993 reforms. In particular, in rich cities where the electoral reform was accompanied by a larger increase in autonomous resources, we document an increase in the percentage of elected mayors with experience in top administrative professions (i.e., with higher administrative skills) and less experience in politics (i.e., with lower political skills). According to our baseline estimates, following the 1993 reforms, a city with a one standard deviation higher GDP per capita (i.e., 3730.80 euros)⁴ ended up with a 6-9 % points higher share of mayors from top administrative professions, and with 1 year less of experience in politics on average. We also show that these effects are not driven by pre-treatment differential trends in the selection of politicians across rich and poor municipalities, and we also provide evidence that in rich municipalities the change in the local political class was followed by an improvement in the quality of services, and that this improvement was effectively due to a “selection effect” on local politicians and not to a stronger “disciplining effect”⁵.

While this evidence is clear, its interpretation in a causal way is much less so. Unfortunately, given all that happened in Italy at the beginning of the '90s, establishing causal links between the increased fiscal autonomy and the observed changes in the characteristics of mayors is difficult. However, we take several potential alternative explanations and assess empirically their validity. First, to discuss the role of the electoral rules, we show that the 1993 reforms had different effects across rich and poor cities *only* with respect to administrative/political skills and not with respect to other characteristics of mayors, like education, sex, age and political orientation. This finding is difficult to reconcile *only* with the change in electoral rules. To further reinforce this evidence, we also perform a placebo test on *national* politicians, which takes advantage of a similar electoral reform implemented at the national level in 1993. This reform, which was clearly not accompanied by any change in tax powers, replaced the old proportional system with a first-past-the-post electoral law, which enabled every electoral district to directly elect one MP. The results show that, after 1993, rich areas elected MPs with *more* political skills, while administrative skills were not affected, which is exactly the opposite of what we observe at the municipal level. This different behavior of voters at the national and at the local level supports our story that the effect on mayors was due to the strategic incentives induced by the different levels of fiscal autonomy more than by the new electoral law. Second, we provide evidence that the results are not determined by the political turmoil that

⁴In most of the empirical analysis we use the pre-reform GDP per-capita as a proxy for fiscal autonomy. This enables to avoid endogeneity issues that characterize a more direct measure of fiscal autonomy as the ratio between municipal own revenues and total revenues. The potential endogeneity is due to the powers of the mayors, who can affect fiscal autonomy changing tax rates and tax bases. As described in section 5.3, our results are unchanged if we use the ratio between municipal own revenues and total revenues rather than the pre-reform GDP per-capita.

⁵The terminology is borrowed from Besley and Smart (2007).

occurred in Italy at the beginning of the '90s following the "Clean Hands" trial, which led to important changes in the political system. More in detail, we exploit a referendum to abolish public funding to political parties, to show that the change in the characteristics of mayors is not driven by a potential shift in voters' preferences against the old political class. Additionally, we show that the birth of new political parties – that followed the political turmoil induced by the discovery of corruption in the ruling political parties – did play a role in the selection of different politicians at the local level, but their impact is not enough to explain the observed change in the characteristics of mayors. Third, we also check whether our findings were not driven by other factors that the literature deems to be important in the selection of politicians. The results are robust to controls for political competition (e.g., Shugart et al. 2005; Atkinson et al., 2009; Galasso and Nannicini, 2011), for the supply of different types of politicians at the local level (such as the presence of entrepreneurs and graduates), for the endowment of "social capital" (e.g., Guiso et al., 2011; Guiso et al., 2016), and for the costs of electoral campaigns. Finally, we show that our results hold even limiting the analysis to Northern municipalities, thus they are not just a simple artifact of the well known Italian North-South divide.

Our findings have important implications for the debate on fiscal federalism, that are more extensively discussed in the conclusions. Clearly, not all recipes are adaptable to all circumstances: fiscal decentralization may be a good idea, but it requires appropriate conditions, in particular a sufficient degree of local fiscal autonomy, in order to work.

This study is linked to different lines of research. Beside the fiscal federalism literature, our work is clearly related to the recent strand of research in political economy and political science that focuses on the effects of institutions on the selection of politicians (e.g., Key, 1956; Besley, 2004, 2005, 2006; Ashworth and Bueno de Mesquita, 2008; Besley and Reynal-Querol, 2011; Caselli and Morelli, 2004; Shugart et al. 2005; Dal Bo, Dal Bo, and Di Tella, 2006; Poutvaara and Takalo, 2007; Mattozzi and Merlo, 2008; Ferraz and Finan, 2009; Atkinson et al., 2009; Gehlbach et al., 2010; Beach and Jones, 2016; Gagliarducci et al., 2010; Folke and Rickne, 2016; Daniele and Geys, 2015; Daniele, 2017; Dal Bo' et al., 2017; Hayeon Joo and Lee, 2018). While most of the effort in this literature has been devoted so far to address the relationship between politicians' wage, electoral competition and the "quality" of politicians, the insight is clearly much more general and could be applied to other types of institutions, including decentralization. Our work adds to this literature the idea that – especially at the local level – skills matter; and voters seem to be able to elect the politicians that are more valuable for their constituency.

The idea that the features of the local political class may depend on the fiscal characteristics of the communities is probably not new, but to the best of our knowledge has not been formalized and explicitly tested before. Only Brollo et al. (2013), looking at Brazilian municipalities, suggest that an increase in transfers may be associated with higher levels of corruption, as the result of both a moral hazard effect on incumbents and a negative selection effect on the pool of candidates for mayors. However, the theoretical result on the selection of politicians is due in this paper to the expansion in the local budget induced by higher transfers (that allows for higher equilibrium rents) and by the assumption that corrupt politicians are more attracted by rents than honest ones. In our case, the effect on the selection of politicians occurs even with an unchanged budget and with politicians of different skills having the same preferences for rents. Additionally, the normative conclusions are different. Brollo et al. (2013) suggest that lower transfers *always* lead to a higher quality of politicians and higher voters' welfare, while in our model this *only* occurs in rich communities. Third, they only study changes in transfers, while we also consider the case of an increase in local taxation accompanied by an offsetting reduc-

tion in transfers. This is important, because it is unclear how a reduction of transfers could always lead to an increase in voters' welfare, in particular in poorer communities. Finally, while they are interested in the quality of politicians measured by their education, we consider here their skills and suggest that different skills can be useful or not for voters depending on the characteristics of local communities.

The remainder of the paper is organized as follows. Section 2 presents a theoretical model of selection of local politicians under different local financial conditions that captures our main idea. Section 3 presents in more details the institutional characteristics of the Italian reforms. Section 4 discusses our econometric strategy. Section 5 is devoted to our main results on the characteristics of mayors, and to several robustness tests. Section 6 presents some evidence on the impact on voters' welfare of selecting mayors with different skills. Section 7 concludes. For the sake of space, we collect the results of several empirical robustness tests in Appendix 1, while the proofs of the theoretical propositions are reported in Appendix 2.

2 The Model

The framework is a standard career-concern model of politics, extended to different types of politicians with specialized skills. Thus, consider a 2 period economy, $s = 1, 2$, where s indexes the period, to which we add a self-selection stage $s = 0$ later on. In this economy, at the beginning of the first period, an incumbent politician is in charge in a sub-national government, partly financed with taxes and partly financed with grants from the central government. At the end of the first period an election takes place and either the incumbent or an opponent is elected to rule for the second period⁶. Local politicians only care about collecting as much rents as possible from office. We let R_s indicate the rents appropriated by an incumbent in period s . Politicians come of two types $j = a, p$: a -type is on average better in organizing local services and raising money from local sources (he has more "administrative" or "managerial" skills), while p -type is on average better in raising money from the central government (he has more "political" skills). Our basic point here is that these are quite different skills, which require a different background and specialization, and are more or less useful depending on the context.
7

Local taxes are fixed so that the voter is only interested in the quality/quantity of local public services, that we capture here with a single local public good g_s . The utility of the voter over the two periods is then just:

$$U = g_1 + \lambda E(g_2) \tag{1}$$

where $0 < \lambda < 1$ is the discount rate and expectations in Eq. (1) are taken with respect to the quality of the politician in the second period (see below). In turn, g_s depends on three factors: positively, on the amount of resources invested in financing it and on the ability of the different types of politicians to raise and use these resources; and negatively, on the rents that the politician in charge diverts to his advantage and to the detriment of voters. More specifically, in period 1, when an incumbent politician of type j is in charge, we assume:

⁶The insights of these career concern models extend to multiple periods. See, for instance, Persson and Tabellini (2000).

⁷In the empirical analysis in Section 5, we proxy these skills by looking at the job of elected mayors before entering in the political arena and at their previous experience in politics. Notice that these information are typically made available to voters at the moment of the elections.

$$g_1^j = [\alpha t \theta^j + \tau(\alpha) \delta^j] (1 - r_1^j) \quad (2)$$

where r_1^j is the rate of rents extraction in period 1 by a politician of type j , t is the exogenously given local tax rate ($0 < t < 1$) and α is the municipal tax base, with $\bar{\alpha} > \alpha > \underline{\alpha} > 0$. $\tau(\alpha) > 0$ is the transfer received by the center. Note that we write τ as a function of α , as grants to local governments are rule-based in most countries, and the rule usually implies larger transfers to poorer communities (e.g. $\tau'(\alpha) < 0$)⁸.

Eq. (2) implies that g_1^j not only depends on the tax base of the municipality and the rule-determined transfers, but also on the ability of politician j to raise extra resources and use them efficiently, captured here by the couple (θ^j, δ^j) . More specifically, a politician with higher level of political skills, δ^j , may be more able to get extra resources from the center, for instance, by lobbying the central government to manipulate the grant formula to the advantage of the municipality; or by convincing it to finance directly some components of local expenditure. Similarly, a politician with higher level of administrative skills, θ^j , may be better able in raising local resources out of the given tax base, for example by improving the collection of local taxes; or be better able in using these resources, generating a higher level of public good provision out of the same amount of local resources. Notice from Eq. (2) that we also assume that a politician can divert to his advantage the extra resources that he himself generates. This is intuitive: a politician with high political skills may cash some of the extra transfers he brings home; a politician with high managerial skills (say, an architect) may divert some of the local funds to his private associates, and so on. Finally, for analytical convenience, we do not allow politicians to take different rents from the different sources of financing: the same rate r_1^j applies to both sources.

Both θ^j and δ^j follow an independent uniform distribution function with density ψ and average $\bar{\theta}^j > 0$, $\bar{\delta}^j > 0$, respectively. In keeping with the discussion above, we assume $\bar{\theta}^a > \bar{\delta}^a$, $\bar{\delta}^p > \bar{\theta}^p$, $\bar{\theta}^a > \bar{\theta}^p$, $\bar{\delta}^a < \bar{\delta}^p$. This captures the idea that a -types are "better" on average in producing local services out of local resources, while p -types are "better" on average in raising extra resources from the center, while still allowing the possibility that for any given realization of (θ^j, δ^j) a politician might be uniformly better or worse than average on both skill dimensions. Let $z^j = \alpha t \theta^j + \tau(\alpha) \delta^j$ be the total municipal revenues generated by an incumbent of type j and let $f(z^j)$ be its density function. $f(z^j)$ can be explicitly derived once the relative range of the extremes of the distribution are determined (see Appendix 2). Specifically, in the following we assume:

- **A.1** $\tau(\bar{\alpha}) > \bar{\alpha}t$; $\frac{\tau(\alpha) - \alpha t}{2\psi} > |E(z^a) - E(z^p)|$.

The first part of assumption A.1 suggests that transfers are the most important component of sub-national government funding. While this fits well the situation of sub-national governments in Italy and in many developing countries currently involved in a decentralization process, it is not essential, and qualitatively the same results could be obtained even imposing the opposite condition $\tau(\underline{\alpha}) < \underline{\alpha}t$ ⁹. The second part of A.1 imposes an exogenous bound on the maximal difference between the two types' (expected) efficiency levels for any level of the tax base. This is also not essential for the results, but allows us to simplify the computations and derive below the equilibrium level of rents.

⁸For notational simplicity, we drop the dependence of τ on a when not needed for the argument.

⁹Details are available by the authors on request.

2.1 The political game

We consider the following political game. At the beginning of period 1, the incumbent j chooses r_1^j , knowing his type and the distribution $f(z)$ for both types, but without knowing the realization of z^j . He also does not know the type of the opponent he is going to face at the elections; he only knows that there is a fraction π of a -type politicians in the population (to be endogenized below), and that the opponent is selected randomly from this population just before the elections. After r_1^j has been chosen, z^j (and, therefore, g_1^j) is also realized. At this stage, nature also chooses the identity (hence, the type) of the opponent. The voter then votes observing g_1^j (but neither r_1^j nor the realization of z^j) and the types of the incumbent and the opponent. The voter also knows $E(z^k)$ for both types, $k = a, p$. With the elections, period 1 ends and period 2 begins. Whoever is in charge at the beginning of period 2 chooses again some rent appropriation for period 2. If the incumbent j is confirmed, the realization of z^j in the first period carries over to the second, as both θ^j and δ^j are permanent characteristics of the incumbent. If instead an opponent k is elected, z^k is realized. In both cases, g_2 is then determined and the game ends.

Assuming that, at the time the incumbent j sets r_1^j , he does not know the realization of z^j is standard in "career concern" models (Persson and Tabellini, 2000). This assumption has the advantage of greatly simplifying the analysis, ruling out signalling effects, while still providing electoral incentives to incumbent politicians. It is also not completely implausible; it just means that voters and candidates may have some common expectations on how a specific candidate will perform in office given his general characteristics, but that his true ability will not be revealed until he has been in charge. Notice that this also implies that all politicians of type j , as they are all ex-ante identical, make the same choice of r_1^j in period 1. The assumption that, at the time he sets r_1^j , the incumbent j does not know the type of the opponent also seems very reasonable, as opponents are typically selected only few months before the elections. However, as will become clear as we proceed, relaxing this assumption would not affect much our results¹⁰.

To solve the model, we work backward. In period 2, as there is no future ahead, whoever is in charge takes maximal rents, $R_2^k = \bar{r}z^k$, where $\bar{r} < 1$ is some maximal rent rate. For analytic simplicity, we assume here that maximal rents an incumbent can cash in both periods take some fix values, independently on j and z^j , i.e., $R_2^k = \bar{R} > 0$ for $k = a, p$ and $s = 1, 2$ ¹¹. In the second period, the utility of the voter is then $z^k - \bar{R}$. This implies that the voter is interested in re-electing (or electing) the candidate with the larger realized (or expected) z^k , as this would produce a higher level of g_2^k .

Having solved period 2, let us go back to period 1. At the end of this period, the voter observes g_1^j but she does not observe either r_1^j or the realization of z^j . The rational voter, however, expects the incumbent to take some rents in the first period. We then look for an equilibrium where the voter uses these expectations to discriminate between high/low quality incumbents. Let r_1^{je} be the rate of rents that the voter expects a politician of type j to take in period 1. Upon observing g_1^j , the expected value of z^j for the voter is then just:

¹⁰It would just mean that at the equilibrium the incumbent would now select a different (expected) level of g_1^j depending of the type of the opponent (see below).

¹¹Assuming $R_2^k = \bar{r}z^k$ would complicate the algebra considerably, without offering extra insights. $R_2^k = \bar{r}z^k$ implies that the more efficient type has even more incentives to refrain from taking maximal rents in the first period, as his expected rents in the second, if elected, are larger. Details are available from the authors on request.

$$E(z^j|g_1^j) = \frac{g_1^j}{(1 - r_1^{je})} \quad (3)$$

Intuitively, the optimal strategy for the voter is then to vote for the incumbent if $E(z^j|g_1^j) \geq E(z^k)$ and vote for the opponent k otherwise. At the equilibrium, the incumbent knows the voter's optimal strategy when setting r_1^j , and knows r_1^{je} . Ex ante, he can then compute the probability of being reelected as a function of r_1^j, r_1^{je} , and of the expected type of the opponent. Using Eq. (2) and (3), and the timing assumptions above, the expected rents of an incumbent of type j over the two periods can be written as:

$$\begin{aligned} E(R^j) = & r_1^j E(z^j) + \lambda \bar{R} \pi \left[1 - \text{prob} \left(z^j \leq E(z^a) \frac{(1 - r_1^{je})}{(1 - r_1^j)} \right) \right] + \\ & + \lambda \bar{R} (1 - \pi) \left[1 - \text{prob} \left(z^j \leq E(z^p) \frac{(1 - r_1^{je})}{(1 - r_1^j)} \right) \right] \end{aligned} \quad (4)$$

Clearly, raising r_1^j increases expected rents in the first period, but for given r_1^{je} , it also reduces the probability of being re-elected in the second, and thus the expected second period rents. At the equilibrium rents rate, the incumbent trades off optimally these two effects. Notice also that, at the equilibrium, voters' expectations need be confirmed, so we look for a solution of the incumbent's problem where $r_1^j = r_1^{je}$ also holds. Invoking Eq. (10) and A.1, deriving and imposing the equilibrium condition $r_1^j = r_1^{je}$, we get r_1^{j*} , the equilibrium rents rate¹²:

$$r_1^{j*} = r_1^{je} = 1 - \frac{\lambda \bar{R} \psi}{\tau} \left[\frac{E(z^o)}{E(z^j)} \right] \quad (5)$$

where $E(z^o) = \pi E(z^a) + (1 - \pi) E(z^p)$. Notice that for r_1^{j*} to be the optimal strategy for both types of incumbent it must also be the case that even the weakest one prefer to play this strategy rather than deviating and taking maximal rents \bar{R} in the first period (and not be re-elected in the second); as shown in Appendix 2, this translates in an upper limit for \bar{R} , $\bar{R} < R'$; in turn, this condition also implies $r_1^{j*} < 1$. At r_1^{j*} , expected rents in the first period for the j 's incumbent are $r_1^{j*} E(z^j) = E(z^j) - \frac{\lambda \bar{R} \psi}{\tau} E(z^o)$; hence, the (expected) level of the public good in the first period is just $g_1^* = \frac{\lambda \bar{R} \psi}{\tau} E(z^o)$ under both types of incumbent. It follows that the more ex ante efficient incumbent, i.e., the incumbent with higher $E(z)$, expects to earn higher first period rents at the equilibrium¹³. Note also that, at the equilibrium, a candidate j expects to be re-elected with probability $\frac{1}{2}$ if he meets at the elections a candidate of the same type, and to be re-elected with probability $\frac{1}{2} + \frac{\psi}{\tau} [E(z^j) - E(z^k)]$, $j, k = a, p, j \neq k$ otherwise. Thus, at the proposed equilibrium, more ex ante efficient types also expect to be re-elected, and earn second period rents, with higher probability.

The strategies of the voter at the proposed equilibrium are straightforward. The voter sets

¹²See Appendix 2 for a formal proof.

¹³As shown, expected rents are instead decreasing in $\lambda \bar{R}$ (a larger $\lambda \bar{R}$ means that second period rents are either larger or that they matter more for the politician, and therefore he is willing to give up more current rents in order to be re-elected) and in the density $\frac{\psi}{\tau}$ (a larger $\frac{\psi}{\tau}$ means that the incumbent expects to lose more votes if r_1^j diverges from r_1^{je}).

up a threshold level for the public good that depends on the type of the incumbent j and the type of the opponent k , $g_1^{jk} = g_1^* \frac{E(z^k)}{E(z^j)}$ and re-elects the incumbent iff $g_1 \geq g_1^{jk}$. Thus, if two candidates of the same type compete in elections, the voter re-elects the incumbent j only if he receives at least $g_1 \geq g_1^*$ in the first period. If instead two candidates of different types compete at the elections, and say, the incumbent of type j is known to be more efficient in expected terms than the opponent of type k , the voter is willing to re-elect j even if he observes a g_1 smaller than g_1^* , provided that g_1 is larger than $g_1^* \frac{E(z^k)}{E(z^j)}$, as this means that, at the equilibrium strategies, the realization of z^j has been higher than the expected value for the opponent, $E(z^k)$. Our results are then similar to the standard ones derived in this literature (Persson and Tabellini, 2000). The important difference is that in our case the voter sets up different thresholds for the public good in the first period, as candidates come of two types and can in turn meet two different types at the elections.

Collecting these observations, we can then conclude:

Proposition 1 *Assume $R' > \bar{R}$. Then there exists a unique equilibrium where the voter sets up a threshold for g_1 , $g_1^{jk} = \frac{\lambda \bar{R} \psi}{\tau} E(z^o) \frac{E(z^k)}{E(z^j)}$ such that she re-elects the incumbent j , if $g_1 \geq g_1^{jk}$, and she elects the opponent k otherwise (where $j, k = a, p$). At this equilibrium, an incumbent j sets first period rents at the rate $r_1^{j*} = 1 - \frac{\lambda \bar{R} \psi}{\tau} \left[\frac{E(z^o)}{E(z^j)} \right]$, where $0 < r_1^{j*} < 1$. $E(z^j) \geq E(z^k)$ ($E(z^j) \leq E(z^k)$) implies that incumbent j weakly earns more (less) expected rents in the first period and he is elected with higher (lower) probability in the second than an incumbent of type k , $j \neq k$.*

Proof. See Appendix 2. ■

2.2 Comparative statics: changes in fiscal autonomy

We now use our results above to study the effects of a tax decentralization reform on the selection of politicians with different skills and consumers' welfare in sub-national governments characterized by a different tax base α . In the context of our model, a tax decentralization reform can be captured as an increase in t and a reduction in τ . Since we are mostly interested in the role of fiscal autonomy, to avoid mechanical effects due to changes in the total resources we consider the possibility for sub-national governments to increase local taxes to be accompanied by an offsetting variation in grants from the centre.¹⁴ In terms of our model, the simplest way to capture this invariance in total resources is by normalizing municipal revenues to unity, hence $\tau(\alpha) = 1 - \alpha t$. It follows that a small increase in t , from t to $t + dt$, in a municipality with tax base α would also imply a reduction in the transfer, $d\tau = -\alpha dt$; and the effect would be different between rich and poor jurisdictions in terms of their tax base. It also follows that $E(z^a) \geq (<) E(z^p)$ for $\alpha \geq (<) \alpha^*$ where $\alpha^* = \frac{\bar{\delta}^p - \bar{\delta}^a}{t(\bar{\theta}^a - \bar{\theta}^p + \bar{\delta}^p - \bar{\delta}^a)}$, provided that $\bar{a} > \alpha^* > \underline{a}$. For concreteness, we assume this to be the case, so that even before the reform there are municipalities (the richest ones) where a -types are in expected terms more efficient than p -types, and other municipalities (the poorest ones) where the opposite is true.

¹⁴Notice that the introduction of the new property tax for Italian municipalities in 1993 was accompanied by an offsetting variation in grants (see Section 3), so that – at the statutory tax rate level of the new tax – each municipality had exactly the same resources both before and after the reform. Thus, financial autonomy was increased in each municipality, but to a different extent depending on the size of the municipal tax base.

What would then be the effect of introducing a compensated tax reform in the context of our model? Consider first the expected welfare of the two types of politicians¹⁵.

Proposition 2 (i) $\partial(E(R^a) - E(R^p))/\partial t > 0$; $\partial^2(E(R^a) - E(R^p))/\partial\alpha\partial t > 0$; (ii) $\partial E(R^p)/\partial t < 0$; $\partial^2 E(R^p)/\partial\alpha\partial t < 0$. (iii) Suppose $\bar{\theta}^a \geq \theta^* > 0$; then $\partial E(R^a)/\partial t > 0$. Suppose $\bar{\theta}^a \geq \theta^{**} > \theta^*$, then $\partial^2 E(R^a)/\partial\alpha\partial t > 0$.

Proof. See Appendix 2. ■

To provide an intuition for Proposition 2, note that a change in t , matched by a revenues offsetting change in τ , has two effects on the expected rents of the two types of politicians. The direct effect is due to the change in $E(z^j)$. Under our assumptions on $\bar{\theta}^j$ and $\bar{\delta}^j$, this effect is certainly positive for the a -type and certainly negative for the p -type. Notice that this also implies that an a -type politician is also more likely to be re-elected in the second period (when meeting an opponent of a different type) as $(E(z^a) - E(z^p))$ also increases. But there is also an indirect effect: the change in t (and, therefore, in τ) increases the density around the equilibrium, $-\frac{\psi}{\tau^2} d\tau/dt > 0$ (see Eq. 5), and therefore reduces expected first period rents for both types of incumbent. Thus, the p -type incumbent is certainly made worse off by the reform. As for the a -type, the total effect depends on the combination of the two effects; and it might be positive if the direct effect dominates the reduction in first period rents. This in turn boils down to this type being efficient enough, that is, on $\bar{\theta}^a$ being larger of some threshold, θ^* . But the important point, also stated in Proposition 2, is that regardless of its effects on the absolute level of politicians' utilities, the tax decentralization reform certainly makes the a -type better off relatively to the p -type, and particularly so in richer communities. This will be useful below.

What about the voters? In the first period, their welfare certainly increases as expected rents for both types fall. But in the second period, signing the effect of the reform is complicated as it clearly depends on the type of the incumbent, the share of the two types of politicians in the population and on the tax base of the communities. To gain insights, it is then more useful to raise the question in expected terms, with expectations taken with respect to the type of incumbents that the voter could face. Let then $U(\alpha) = \pi U^a(\alpha) + (1 - \pi)U^p(\alpha)$ be the expected utility of a voter living in a local government with tax base α , where $U^j(\alpha)$ is consumers' expected utility over the two periods when the first period incumbent is of type j , $j = a, p$. One can then show the following:

Proposition 3 Assume $\bar{\theta}^a > \bar{\theta}^{a*} > \bar{\theta}^{p*} > \bar{\theta}^p > 0$. Then, (i) $\partial U(\alpha)/\partial t < 0$ for $\pi \rightarrow 0$; $\partial U(\alpha)/\partial t > 0$ for $\pi \rightarrow 1$. (ii) There exists a unique value $\underline{\pi}(\alpha) > 0$ such that $\partial U(\alpha)/\partial t = 0$, and $\partial U(\alpha)/\partial t > (<)0$ for $\pi > (<)\underline{\pi}(\alpha)$. (iii) $\partial \underline{\pi}(\alpha)/\partial \alpha < 0$.

Proof. See Appendix 2. ■

Thus, quite intuitively, whether the voter benefits or is damaged by the reform depends on the share of a -type politicians, and on the tax base of the community where she lives. In particular, provided that the polarization in skills between the two types of politicians is large enough, and in spite of the negative effect on first period rents, the voter is certainly made worse off by the reform if all politicians are of p -type and certainly made better off if all politicians

¹⁵In the Propositions to follow, when we differentiate for t we take into account the dependence of τ on t , that is $\tau(\alpha) = 1 - \alpha t$.

are of a -type. This holds irrespective of the tax base, although in the latter case consumers in rich jurisdictions gain the most from the reform. However, for intermediate values of π , it is the tax base that matters in determining the welfare effect of the reform. In particular, for given π , the richer is the community, the more likely it is that the consumer benefits from the reform.

Thus, our model certainly does not support the claim that consumers *always* benefit by a tax decentralization reform, even if the reform is compensated by an offsetting variation in transfers. Intuitively, in poor communities, the reform just reduces the usefulness of the p -type politicians to voters, still maintaining them as the more efficient politicians, and therefore the ones more likely to be elected.

2.3 Endogenous candidacy

So far we took π as given. But as the change in financing rules also changes the expected rents for both types for entering in politics, one would expect that the reform also affects both the size and the composition of the set of candidates. To study this case, suppose that at time 0, before period 1 begins, a citizen of type j , $j = a, p$, is considering whether entering in the political arena. Suppose that there are n such potential candidates, where n is assumed to be a quite large number. The candidacy choice depends on the opportunity cost for entering in politics, that is, on the remuneration that a potential candidate of type j could alternatively earn if she decided to remain a private citizen instead. Let us assume that the wage that each of the potential candidates j earns in the private market is drawn at the beginning of period 0 from a common independent uniform distribution on the interval $\{0, \bar{w}\}$. Citizen j observes the realization of her wage w^j before deciding whether becoming a member of the set of potential candidates j ; she also knows the expected two period rents for becoming an incumbent at time 1, $E(R^j)$. We assume that there are no costs in joining the set of potential candidates, and that both w^j and $E(R^j)$ are so large with respect to the benefits/costs that j receives from the municipality as a private citizen that she just ignores the latter in taking her candidacy decision. The only cost for a citizen j of becoming a politician is that if she is elected, she has to rule, giving up her private wage. The candidacy decision is taken at the end of time 0 and cannot be revised afterward. After the candidacy choice has been taken and so the set of all possible politicians at the end of period 0 is determined, one candidate is chosen randomly by nature to become the incumbent politician in charge at period 1; the game then unfolds as already described in the previous Section.

Under these assumptions, the choice of citizen j at the end of time 0 is quite simple. She will accept to join the set of politicians if the expected rents from doing so (in the case she is selected to become the incumbent politician at period 1) overcome the foregone wages; that is, provided $E(R^j) \geq (1 + \lambda)w^j$. The ex-ante probability (computed at the beginning of time 0, before the realization of w^j) that a citizen of type j joins the political market is then $\frac{1}{\bar{w}} \frac{E(R^j)}{(1+\lambda)}$, and as all j face the same distribution, the expected number of individuals of type j (equal to the realized number for large n) who join the political market is then $J = \frac{1}{\bar{w}} \frac{E(R^j)}{(1+\lambda)} n$, where $J = A, P$. It immediately follows that $\pi = \frac{A}{A+P} = \frac{E(R^a)}{E(R^a)+E(R^p)}$.

Notice from the discussion above that while π depends on $E(R^j)$, $E(R^j)$ also depends on π , as the probability of being re-elected (and, therefore, second period rents of an incumbent) depends on the probability of meeting different types of opponents. Intuitively, higher expected rents for the more efficient type j induce more individuals of type j to enter the political market which in turn reduces expected rents, as it reduces the probability of meeting an opponent of

a less efficient type. At the equilibrium π^* these two forces need to balance. As Appendix 2 shows, solving the resulting system of simultaneous equations and assuming an interior solution, this equilibrium share can be computed as:

$$\pi^* = \frac{1}{2} + \frac{(E(z^a) - E(z^p))(1 + \frac{\lambda \bar{R}\psi}{\tau})}{c(A + P)} \quad (6)$$

where $c = \frac{2\bar{w}(1+\lambda)}{n} > 0$. Thus, the denominator of the second term on the RHS of Eq. (6) is just proportional to the total number of potential politicians of both types and it is strictly positive. Eq. (6) allows us to get an important conclusion:

Corollary 1 $\alpha \geq (<) \alpha^*$ implies $\pi^* \geq (<) \frac{1}{2}$.

Hence, at the equilibrium with endogenous candidacy, richer municipalities have a larger share of a -type politicians and, therefore, in expected terms, a higher share of incumbent politicians of a -type¹⁶. Using Eq. (6), we can also investigate the effect of the reform on π^* :

Proposition 4 $\partial\pi^*/\partial t > 0$ for $\pi^* \geq \frac{1}{2}$. $\partial\pi^*/\partial t > 0$ for π^* smaller but close to $\frac{1}{2}$. For lower values of π^* , the sign of $\partial\pi^*/\partial t$ is uncertain and might become negative.

Proof. See Appendix 2. ■

Thus, considering endogenous candidacy, the reform should have the effect of increasing even further the divergence between jurisdictions. After the reform, richer communities should have even more a -type politicians, while poorer communities would have a much lower increase or indeed a reduction in π^* . The intuition is simple. A revenue compensated increase in t would certainly have the effect of increasing the numerator of the second term on the RHS of Eq. (6), as a -type incumbents become relatively more efficient than p -types. Under a very mild condition, discussed in Appendix 2, the same reform would also have the effect of reducing the total number of politicians (the denominator in Eq. 6), as many p -type politicians would leave the political market and a -types may also leave the market (if $E(R^a)$ falls following the reform, see Proposition 2) or in any case the increase in their number is not enough to compensate for the exit of the p -types. Where $\pi^* \geq \frac{1}{2}$, the two effects work in the same direction, thus leading to an increase in π^* . Where $\pi^* < \frac{1}{2}$, but not too far from $\frac{1}{2}$, the first effect still dominates the second, so leading to an increase in π^* albeit at a reduced rate. Finally, for poor communities, the second effect may dominate, leading to a further reduction in π^* .

As for voters' welfare, endogenous candidacy clearly emphasizes previous results. Proposition 3 suggests that a compensated tax reform, increasing fiscal autonomy at unchanged total resources, is more likely to benefit the voters the richer is the community and the greater is the share of a -type politicians. Proposition 4 suggests that this reform should also increase the share of a -type politicians more (or only) in the richer jurisdictions. Hence, in rich communities the tax reform should increase the expected welfare of voters for two reasons: directly, because it increases the utility of voters for given π ; and indirectly, because it also increases π . On the

¹⁶Private market opportunities could of course differ among types. For instance, it might be that the market opportunities for the a -type are larger in richer communities, implying $\bar{w}^a > \bar{w}^p$. Adding this complication to the model would forbid us from getting an explicit analytic solution for π^* . But it would not affect the comparative static results below. Details available on request.

contrary, in poor communities, the reform might decrease voters' welfare directly, as the more efficient p -type becomes less useful to voters, and indirectly, as π does not increase much, or even decreases.

Summing up, our model then produces the following predictions. A tax decentralization reform which increases tax autonomy for sub-national governments should have the effect of increasing the share of politicians with high administrative skills in rich communities. In poorer ones, the increase of the share of this type of politicians is smaller and might even become negative. The effect on voter's welfare also depends on the relative wealth of the communities. Consumers' welfare should increase in rich jurisdictions, while the effect on poor ones is uncertain and might even be negative. Bearing these predictions in mind, let us then discuss our case study.

3 A case-study: the Italian decentralization reforms

We look for supportive evidence to our argument by considering the Italian experience of tax decentralization involving municipalities. The beginning of the '90s was a period of turmoil in Italy, on both economic and political grounds. A financial crisis in 1992 led the Italian lira, together with the British pound, to abandon the European exchange rate agreements and forced the governments of the time to launch a severe fiscal consolidation policy that lasted up to 1997, when the country eventually obtained the access to the Euro area meeting Maastricht requirements. An investigation led by judges in Milan (the so-called "Clean Hands" trials), beginning in 1992, proved the extent of corruption in the ruling political parties, and increased the demand by citizens for more responsiveness and accountability at the political level. Under the pressure of public opinion, who also directly expressed itself in national popular referenda held on April 1993, electoral rules at all levels of government were changed, abandoning the traditional proportional electoral rule in favor of systems that made elected politicians more directly accountable to citizens¹⁷. The new electoral rules and the loss of credibility induced by corruption charges, emphasized by media, also led to the disappearance of the political parties that had continuously ruled the country in the aftermath of World War II and to the birth of new ones (opening the doors to the so-called "Second Republic"), such as *Forza Italia*, the party of Mr. Berlusconi, and *Lega Nord*. Finally, the need to improve financial stability at the sub-national level convinced the central government of the need to increase the extent of autonomous financing at local level, reforming a system that was traditionally based on transfers and tied grants¹⁸. This began what is known as the "decentralization" period in Italy, that culminated in the 2001 Constitutional reform (see Ambrosanio et al., 2015).

For municipalities, an important political body in the Italian architecture of governments¹⁹,

¹⁷Specifically, electoral rules at the national level were changed for both houses in August 1993 (Laws 276 and 277), at the municipal and provincial level in March 1993 (Law 81) and at the regional level in February 1995 (Law 43). For a thorough discussion on the Italian electoral systems, see D'Alimonte (2001) and Katz (2001).

¹⁸Specifically, regions received new tax tools in 1995 and 1998, and provinces in 1998. In all cases, grants were reduced proportionally, following the example of the property tax for municipalities, discussed below in the text. The allocation of expenditure functions remained instead unchanged up to 1998, when extra functions were devolved to regions and provinces. See Ambrosanio et al. (2015).

¹⁹Differences among the about 8,000 Italian municipalities are huge, with respect to size, average income, population density and composition. They are in charge of a large number of services, ranging from general administrative services (like the registry office) to local public good provisions (like local transports, public parks and amenities, street lighting and cleaning, urban and sport infrastructures, maintenance of school's building,

this resulted in the passing of two fundamental reforms. The first concerned the electoral system. In 1993 (Law 81/1993), the traditional parliamentary system was substituted by a quasi-presidential one, involving the direct election of mayors. Before the reform, citizens voted for parties' representatives to elect the city council, that then elected the mayor and the executive office. Since the reform, citizens directly elect the mayor, and a majoritarian prize guarantees that parties supporting the winning candidate also get the majority of the seats in the city council²⁰. The mayor is then free to choose (and dismiss) her executive office. The city council can still dismiss the mayor, but in that case new elections need to take place. The reform also introduced a term limit for the mayor, that could now be elected for two consecutive rounds only²¹. Finally, for further reference, notice that the mayor's wage, as well as the number of councilors and their salary, are determined uniformly across the country according to national rules²². All these variables are positively related to the population of the municipality, but do not depend on local revenues (e.g., Gagliarducci and Nannicini, 2013). As pointed out by a large literature in political sciences (e.g., Pasquino, 2006; Baldini and Legnante, 2000; Bettin Lattes and Magnier, 1995), it is clear that - after the reform - Italian mayors assumed a paramount role in municipal policy. The reform also changed the nature of local politics, emphasizing the role of the personal characteristics of mayors.

Second, in the very same year, 1993, a fiscal reform introduced a new property tax (ICI) for municipalities, based on the value of all buildings and lands (Legislative Decree 504/1992), providing Italian cities with a large autonomous source of tax revenues. The tax base was determined uniformly across the country (using the national Cadastre), but municipalities were given some autonomy in the setting up of tax rates and tax allowances²³. The central government compensated revenues from the newly introduced tax with a reduction in grants, so that at the minimum (compulsory) ICI tax rate, each municipality's revenues remained unchanged.

Finally, in 1999, a further reform, at unchanged grants and electoral rules, offered an additional autonomous source of funding to municipalities, allowing them to impose a surcharge on the Personal Income Tax of residents, up to 0.5% of the statutory rates of the national Personal Income Tax (PIT).

As shown in Figure 1, these funding reforms had a dramatic effect on the composition of municipal revenues. For Italian municipalities as a whole, the share of transfers out of total revenues fell from just above 60% in 1992 to about 40% in 1993. This share fell further in 2000, albeit to a lower extent, following the introduction of the PIT surcharge. More important, the composition of municipal revenues changed differently across municipalities, depending on the distribution of the tax base: in 2000, for instance, municipalities in the richer North were *on average* financed with own revenues for above 70% of their budget, while in the poorer South

kindergartens), from environmental services (garbage collection and disposal) to public utilities (heating and water provision).

²⁰There is a difference in the electoral rules between municipalities with less than 15,000 inhabitants (that elect the mayor in a single ballot), and municipalities above 15,000 inhabitants (that instead use a run-off). These differences are explained and exploited in Bordignon et al. (2016). Note, however, that all municipalities considered in our empirical analysis are above the 15,000 threshold, so that the electoral rules are the same.

²¹The duration of the municipal legislature was 5 years before 1993, it was reduced to 4 years with the reform, and brought back to 5 years in 2000. We exploit these institutional characteristics in the analysis of the quality of services below.

²²In ordinary statute regions, which is why we only focus on these regions in the subsequent analysis.

²³Tax rates could be set in an interval between 0.4% and 0.7%, differentiating the rates between residential housing and commercial buildings. Municipalities could also introduce an allowance for resident house owners. See Bordignon et al. (2003) and Bordignon et al. (2017) for further details.

grants covered *on average* about 60-70% of total municipal expenditure²⁴.

4 The empirical strategy

The Italian case offers an interesting testing grounds for our theory. First, the electoral reform allowed citizens to choose directly a mayor with specific characteristics. Second, as a result of the decentralization reform, fiscal autonomy also changed dramatically at the municipal level, both over time and across municipalities, making administrative/political skills more/less useful characteristics of mayors in different cities. According to our theory, this should have led to more mayors with administrative skills and less political skills in richer cities with respect to poorer ones.²⁵ In turn, this should have influenced voters' welfare, via for instance a more efficient collection of taxes or improved quality of services. In the next two sections, we then check if indeed after the tax reforms richer cities elected mayors with more administrative skills and less political skills with respect to poor cities; and whether we also observe a positive impact on municipal services.

We begin by adopting a simple difference-in-differences approach (e.g., Angrist and Pischke, 2009), based on the idea that municipalities were "treated" by the 1993 fiscal decentralization reform. In this case, the dummy variable $Decentr_t$ identifies the post-treatment years (i.e., all the years following the introduction of the ICI property tax²⁶), while $FiscAut_i$ is our measure of fiscal autonomy (see the definition below). The interaction term $Decentr_t \times FiscAut_i$ represents our treatment variable, allowing the impact of the tax decentralization reform to differ across rich and poor cities²⁷. Our implicit assumption here is that while the shock on electoral rules should have impacted similarly across the country, the shock on fiscal autonomy would have produced different results in different municipalities depending on their tax base. This basic model can then be written as:

$$Y_{it}^j = A_i + B_t + \gamma X_{it} + \delta Decentr_t \times FiscAut_i + \varepsilon_{it} \quad (7)$$

where the superscript $j = a, p, w$ identify three different outcome variables of interest (namely, a -type politicians, p -type politicians and consumers' welfare, respectively, to be defined below in Section 5 and 6); A_i and B_t are fixed effects for municipalities and years; X_{it} is a vector of political and economic variables that capture observable differences across municipalities. Estimates of δ offer us preliminary evidence of the heterogeneous impact of tax decentralization reforms on outcomes. In particular, on the basis of our theoretical model, we expect δ to be positive for both Y^a and Y^w , and negative for Y^p .

²⁴See ISTAT, *I Bilanci consuntivi delle amministrazioni comunali*, June 2012, available at www.istat.it. Unfortunately, ISTAT does not provide a regional classification for all municipalities until 2000.

²⁵This is due to citizens preferring to vote for candidates with administrative rather than political skills, as the former skills were now more useful to them. Voters' behavior could also be enhanced by a self-selection effect, as after the reform more candidates with administrative skills would run for mayors in richer cities. As explained below, we cannot distinguish empirically between these two effects because before the electoral reform there were no candidates for mayors. Under the parliamentary system in place before the reform, mayors were selected by the winning political parties among the elected municipal counselors after the elections, and often changed in between elections.

²⁶We use the first year an election took place after the introduction of the property tax (ICI) as the first year of the treatment, as both the mayor and the municipal council are renewed only after an election. In our sample, more than 2/3 of municipalities held their first election in the years after 1993.

²⁷In this sense, we follow a multivalued treatment approach; see, e.g., Imbens and Wooldridge, 2009.

We then extend this simple model in two different directions. First, we separate the role of the two consecutive tax reforms experienced by the Italian municipalities. In particular, we define two different post-treatment periods: $Decentr1_t$ identifies the years between the introduction of ICI and the introduction of the PIT surcharge (1993-1998), while $Decentr2_t$ takes value one after the introduction of the PIT surcharge (1999-2002)²⁸. The important feature of this second approach is that, while $Decentr1_t$ identifies the shock stemming from the new electoral rules and the new property tax, $Decentr2_t$ identifies the *additional* shock on fiscal autonomy at *unchanged electoral rules*, so that any further differential effect stemming from the treatment variable $FiscAut_i$ could be attributed solely to tax decentralization in this second period. This offers a first attempt to separate the impact of tax decentralization from the impact of the new electoral rules. The second model can be written as:

$$Y_{it}^j = A_i + B_t + \gamma X_{it} + \delta_1 Decentr1_t \times FiscAut_i + \delta_2 Decentr2_t \times FiscAut_i + \varepsilon_{it} \quad (8)$$

The coefficients of interest here are δ_1 and δ_2 , that capture the differential impact on political selection of the ICI property tax (δ_1) and the additional contribution coming from the PIT surcharge (δ_2). Clearly, we expect δ_1 and δ_2 to be positive for both Y^a and Y^w and negative for Y^p .

Second, given all that happened in Italy in the same years, we also control for several other factors that might have affected the selection of local politicians, by defining additional interactions with our dummy $Decentr$:

$$Y_{it}^j = A_i + B_t + \gamma X_{it} + \delta Decentr_t \times FiscAut_i + \theta Decentr_t \times X_i + \varepsilon_{it} \quad (9)$$

If our story is correct, we would still expect δ to be positive (and statistically significant) for both Y^a and Y^w , and negative for Y^p even after controlling for these additional interactions.

5 Evidence on the skills of elected mayors

5.1 Data and variables definition

Our empirical analysis requires variables that capture the characteristics of both the mayors and the municipalities. Data on politicians' characteristics are provided by the Italian Ministry of Domestic Affairs²⁹. The data-set includes information on sex, age, date and place of birth, party affiliation, level of education, and – importantly – the profession of the mayor before entering politics. For the variables capturing municipal characteristics, we exploit different sources commonly used in the empirical works in this area, which are described in Appendix 1, Table A.1.

²⁸Our analysis stops in 2002 because the PIT surcharge was frozen by the central government in that year, to be resumed only in 2007. Moreover, the Central government eliminated the property tax on the main residences of taxpayers in 2008 and reintroduced it again in 2012 with very different characteristics.

²⁹Anagrafe degli Amministratori Locali e Regionali: <http://amministratori.interno.it/>.

The analysis is based on data relative to the 89 provincial capitals (*Capoluogo di Provincia*) of all the ordinary statute regions for the years 1988-2002, a period that includes all the reforms described above. We focus on provincial capitals for different reasons. First, these are all relatively large municipalities subject to the same voting rules (runoff elections) and the same fiscal rules (the Internal Stability Pact). Second, these are municipalities for which we can track back the career of elected mayors from various sources, such as internet web pages and newspapers' archives, in order to build our measures of administrative and political skills. Third, data on the quality of services, in particular on the percentage of "separate" waste collection, are available only for the provincial capitals.

Unfortunately, institutional rules in place before 1993 do not allow us to study to what extent the effect of the decentralization reforms is due to self-selection of candidates and to what extent to voters that select a specific mayoral candidate. In fact, given the proportional electoral system in place before 1993, political parties were not indicating mayoral candidates before the election, and the mayor was selected only after the election by the winning coalition. These rules justify our focus on elected mayors only.

Finally, in the empirical analysis we use all the 15 years in the period 1988-2002, including non-electoral years, to account for the fact that before the electoral reform, mayors could be, and often were, changed during the same electoral mandate between two elections. However, to address the concern that this might affect the efficiency of our estimates, in a robustness exercise we drop all non-electoral years and re-estimate equation (7) using only data for electoral years (see section 5.3).

5.1.1 Outcome variables: skills

A difficult task for our empirical analysis is to define appropriate measures for "administrative" and "political" skills of mayors. As for the former, to identify good administrators we exploit information about *the profession of the mayor before entering politics*. We consider different definitions. First, we use an index – built by the Italian Institute for the Professional Training of Workers (ISFOL), a public research institute supervised by the Italian Ministry of Labour and Social Policy – measuring how important "administrative" skills are for each specific job. The index is defined on a 0-100 continuous scale, and it is built starting from a survey in which a sample of workers is asked to rate (from 0 to 100) how important is a certain skill for their job, and how complex is a certain task for which this skill is required. We consider in particular the skill "manage financial resources" to pick up *a*-types. Therefore, the variable *adm_index*, our first definition of Y^a , maps each job held by the mayor before entering in politics with the ISFOL index, providing a continuous measure of the managerial abilities of elected mayors. Second, we also create two dummy variables that classify the professions characterized by high levels of administrative skills. The first dummy variable (*adm_jobs[A]*) is equal to one for jobs with a value of the ISFOL index greater than the median of the distribution. This includes school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyors, business consultants and pharmacists. The second dummy variable (*adm_jobs[B]*) is defined, somewhat more arbitrarily, by excluding from *adm_jobs[A]* a number of professions whose level of administrative skills is more disputable³⁰. Thus, in this final and more restricted definition, we only remain

³⁰For instance for certain jobs, like for biologists, the index appears to be disproportionately high, since it is affected by the presence of directors in research institutes or hospitals. Given that it is not possible to separate biologists or veterinarians which are also directors from the ones that are not, we decided to exclude these

with school managers, entrepreneurs, directors, engineers and architects, traders, surveyors and business consultants.

As for the p -type politicians (Y^p), we proxy “political” skills with the *previous political experience of the candidate at all levels* before becoming mayor. We use several sources to identify all the political offices previously held by an elected mayor, including all legislative and executive positions in all levels of local governments (regions, provinces, municipalities) and the Italian and the European Parliament.³¹ We then define the variable $\%pol_exp$ as the ratio between the years of previous political experience and the working age of the mayor (i.e., his/her age minus 17). Hence, $\%pol_exp$ is a continuous variable measuring the percentage of working years a candidate had already dedicated to politics before becoming mayor. This should help us in identifying professional politicians, that is individuals whose main job is in the political arena. However, as political experience in itself may be important, we use as a second measure for Y^p the numerator of the variable $\%pol_exp$, i.e., the number of years of previous political experience ($years_pol_exp$) regardless of the age of mayor.

Descriptive statistics for our outcome variables are in Table A.1 in the Appendix. The continuous index measuring the importance of administrative skills has a mean of 51.21 on a 0-100 scale, with a range between 27 and 75.75. According to our two definitions of jobs characterized by administrative skills, this translates in 21-29 percent of elected mayors who are a -types on average depending on the definition used. As for political skills, mayors in our sample spent 26 percent of their career in politics, on average, with an observed range between 0 (never been in politics before) and 93 percent (an entire career devoted to politics). Average years of political experience are about 8, with a range from 0 to 56.

5.1.2 The treatment variable

The dummy variables identifying post-treatment years are interacted with the variable $FiscAut$, a measure of fiscal autonomy. We proxy the fiscal capacity of a municipality with per capita GDP in the province³² in which the city is located³³. This is highly correlated with the tax base of both the property tax and the PIT, and it is less subject to endogeneity concerns³⁴. To avoid endogeneity issues even further, we use per-capita GDP averaged over the pre-treatment years 1991-1992. We observe a large variability in our sample: the range of per-capita GDP goes from 8,851 to 26,041 euro³⁵; this variability will be helpful in identifying our coefficients of interest.

professions from the third definition of Y^a . In the same way, we also excluded university professors because of the heterogeneity of this definition. Similar reasoning was applied to other jobs.

³¹We consider in particular: the online registry office of the Italian Ministry of Internal Affairs for all local levels of government; the online historical archive of both the Italian Chamber of Deputies and the Senate of the Italian Republic; the online archive of the European Parliament. We also use the web page of “Openpolis”, and the historical archive of both *Corriere della Sera* and *La Repubblica*, the two most important Italian newspapers.

³²Provinces are administrative and political territorial bodies intermediate between the regions and the municipalities. In the Italian ordinary regions as a whole there are 89 provinces (on average 6 provinces for region) and each of the city in our sample it is the administrative capital of one of this province.

³³Data are provided by the “IstitutoTagliacarne” (<http://www.tagliacarne.it/>). We use GDP measured at provincial level because data at the municipal level are not available for all years.

³⁴However, in a robustness exercise below (see section 5.3) we also experiment with alternative measures of fiscal autonomy at municipal level, getting comparable results.

³⁵In the regressions, to facilitate the readability of coefficients, we divide GDP by 10,000, so that our unit of measure for GDP per capita is in tens of thousands.

5.2 Results

5.2.1 Main results

The baseline evidence. We estimate all our models by OLS, using two-way cluster-robust standard errors to account for both serial and panel correlations (Cameron et al., 2011). Table 1 reports the baseline estimates of our coefficient of interest δ for all the five different definitions of the dependent variables Y^a and Y^p in the simple model (7) where we only control for year and municipal fixed effects. These estimates represent the raw effects of the reform. Results for the three variables relative to administrative skills are reported in Panel A: the (log of the) ISFOL index (*adm_index*) in column (1) and the two dummy variables derived from it (*adm_jobs[A]* and *adm_jobs[B]*) in columns (2)-(3). All the estimated coefficients in Panel A take up the expected positive signs, are statistically significant at the usual confidence levels, and show reasonable magnitudes. In particular, according to our estimates, following the 1993 decentralization reforms, a city with a one standard deviation higher GDP per capita (i.e., 3730.80 euros) witnessed an increase in the ISFOL index (*adm_index*) of approximately 5%, and an increase in the proportion of mayors coming from jobs characterized by high administrative skills between 9% points (*adm_jobs[B]*) and 6% points (*adm_jobs[A]*). In Panel B, we report the results for the proportion of *p*-type politicians. Also in this case, in line with our theoretical predictions, the raw effects in the first two columns of panel B show that richer cities experienced a sharper decrease in the proportion of *p*-type politicians in the aftermath of the 1993 reform. The effect is statistically different from zero and the magnitude of the coefficient is important from an economic point of view. Specifically, considering again a city with a one standard deviation higher GDP per capita, after the electoral reform, the percentage of past working years that mayors had dedicated to politics before being elected (*%pol_exp*) fell by approximately 3.4% points, while average past political experience (*years_pol_exp*) declined by about 1 year.

In Table 2 (columns (1), (3) and (5)) and Table 3 (columns (1) and (2)) we report the results obtained by running the second model (8), which separates the effect of the two tax reforms and allows to test the common trend assumption (see below). As shown, the raw effects in Table 1 are largely confirmed, in terms of sign, magnitude, and statistical significance. Importantly, we obtain a positive and statistical coefficient for *a*-type politicians and a negative one for *p*-type politicians not only after the introduction of the property tax in 1993 (that occurred simultaneously at the change in the electoral rules) but also after the introduction of the PIT surcharge in 1999 (that occurred at *unchanged* electoral rules).

As we will see below, adding covariates to both models does not change these raw effects, while augmenting our baseline specification with additional interaction terms to control for other possible factors affecting the selection of politicians actually increase the magnitude of our coefficients of interest, leaving signs and statistical significance unaffected.

Common trends assumption. Difference-in-differences estimates rely on the assumption of a common trend for all municipalities. We check the validity of the assumption by following two different strategies. First, we add two anticipatory effects to model (8) interacting our measure for fiscal autonomy (*FiscAut_i*) with two dummy variables *Before₋₂* and *Before₋₁*, which take value 1 two years and one year before the introduction of the property tax, respectively. Results for this exercise are reported in columns (2), (4) and (6) of Table 2 for administrative skills, and in columns (2) and (4) of Table 3 for political skills. As shown, the coefficients for anticipatory effects (δ_{-2} and δ_{-1}) are always insignificant, suggesting that there is no evidence of differential

trends in pre-treatment years.

Figures 2 and 3 provides a useful visual representation of our results: the coefficients of the anticipatory effects (δ_{-2} and δ_{-1}) are not statistically different from zero, while as soon as the reform is implemented and elections take place, *a*-type politicians increase, while *p*-type politicians decrease, and the more so in more fiscally autonomous municipalities.

As a second check we also perform a placebo test. We consider pre-treatment years 1988-1992, and we introduce a fake reform, identified by a dummy variable which takes value 1 starting from 1990 (*fake_decentr*). We interact this variable with our measure for fiscal autonomy *FiscAut*, and estimate model (7) for all our outcome variables considering pre-treatment years 1988-1992 only. Our coefficients of interest from this specification are never statistically significant (see Table A.2 in Appendix 1), further confirming the absence of any bias for differential trends.

5.2.2 Discussion: a causal effect? Alternative explanations

Our evidence so far unambiguously show that after the reforms – which gave more importance to the personal characteristics of mayors and increased fiscal autonomy – we did observe a differential change between rich and poor cities in the ruling class at municipal level. In rich cities, elected mayors are now characterized by higher administrative/managerial skills, and less characterized by political skills. Our theoretical argument attributes this evidence to the shock occurred in fiscal autonomy, which made administrative skills more useful in rich jurisdictions. However, the causal interpretation of our findings is subject to a number of concerns. Given all that happened in Italy at the beginning of the '90s, one may well suspect that the differential effect on elected mayors might be due to other factors that also somehow interacted with a city's fiscal capacity. Unfortunately, from an empirical point of view, there are no easy ways out from this problem. Hence, what we do in the following is to carefully review a number of potential alternative explanations. If these alternative stories cannot explain (or cannot entirely explain) the evidence, then we can conclude that our argument did likely play a role in the observed change in the characteristics of elected mayors.

The role of the electoral law As already noticed, one of the main threat to the causal interpretation of our findings is represented by the simultaneity between the introduction of the new electoral law and the new property tax. An electoral law granting voters the possibility of directly choosing the mayor is essential for our argument, since voters should have the opportunity to pick a mayor with certain characteristics depending on the fiscal autonomy of the municipality. But one could easily argue that is the variation in electoral rules *per se* that brought cities with different fiscal autonomy to choose different types of politicians, rather than being a consequence of the different degree of fiscal autonomy across municipalities. The significant coefficient for the interaction between *Decentr2* and *FiscAut* in Table 2 and 3, relative to the introduction of the PIT surcharge at unchanged electoral rules, it is already reassuring that our theory contributes to explain the evidence we observe. However, to further address the problem, we provide here additional exercises.

First, we use as dependent variables additional characteristics of the mayors. If the new electoral rules played a differential role across rich and poor municipalities, then we should observe the selection of different types of politicians not only in terms of administrative/political skills, but also in terms of education, gender, age and political orientation. For instance, one might expect more educated and younger individuals to be more attracted by the higher

visibility granted by the direct election of the mayor in rich municipalities rather than in the poor ones. At the same time, the new electoral law might also have heterogeneously affected the incentives to enter politics of male and female individuals, or of politicians with different political orientations. To check for this, we run model (7) using the following dependent variables: a dummy variable for graduated mayors³⁶; a dummy for female mayors; the variable *age* (i.e., the age of the mayor); and a dummy for centre-left oriented mayors. As shown in Table 4 (and visualized in Figure 4), none of these variables changed differently across rich and poor communities following the 1993 reforms. The fact that no other characteristics of the political class, except the skills discussed by our theory, changed across rich and poor communities is difficult to justify with the change in the electoral law, bringing support to our idea that is the shock to fiscal autonomy that matters.

Second, we consider voters' behavior in a different context, implementing a placebo test on the characteristics of *national* politicians. If the observed effect on mayors was driven by the new electoral rule, then we should observe a similar effect at the national level, as also in this case the old proportional system was substituted in 1993 by a system mirroring the direct election of mayors (i.e., a first-past-the-post system in single-member districts for 75% of the seats in both Houses). The crucial intuition of this placebo exercise is that the reform of the electoral law at the national level was clearly not accompanied by any changes in tax powers comparable to the introduction of the ICI property tax at municipal level. Thus, any change in the characteristics of national MPs can be attributed solely to the change of the electoral law. To implement this placebo test, we use the data-set built by Gagliarducci et al. (2010) to define our proxies for *a*-types and *p*-types national MPs, and then estimate model (7). The data used in this exercise covers five rounds of national elections (1987, 1992, 1994, 1996 and 2001), the first two run with the old proportional electoral law and the last three with the new majoritarian system. We use only data for MPs elected at the Lower House (*Camera dei Deputati*), because for Senators only the region of election is reported. As for the dependent variables, we measure the presence of *a*-types MPs employing the same variables used for mayors (*adm_index*, *adm_jobs[A]*, *adm_jobs[B]*), while for *p*-types we use the following variables contained in the Gagliarducci et al. (2010) data-set: a dummy variable equal to 1 for professional politicians and trade unionists; the number of past terms a politician had served as MP; the number of past years he had served as MP; a dummy variable equal to 1 for MPs elected at other levels of government in the past. Although not exactly coinciding with the political skills variables we used for mayors, these variables capture the political experience of an elected MP. Finally, we also use as dependent variable the proportion of graduated MPs. Results from this exercise are reported in Table A.3 (Appendix 1). The geographical units used are the Italian provinces³⁷. We run the regressions adding to the model provincial and year of election fixed effects. The coefficients for the *a*-type MPs are never statistically different from zero, and are smaller in magnitude compared to the coefficients estimated for municipalities in the baseline exercise in Table 1. On the contrary, for political skills, the coefficients are in general

³⁶Notice that the level of education of a politician, and in particular whether she holds a university degree, has been often used in previous research (e.g., Gagliarducci and Nannicini, 2013 for Italian mayors) as a measure of quality.

³⁷The multi-member constituencies of the old proportional system are wider and in general different from the single-member districts used under the first-past-the-post system. Furthermore, districts generally contain multiple cities, so that it is in general impossible to recover the city of election for national MPs. To overcome these issues, we aggregate the data at provincial level, calculating the average values of the variables used by province.

positive and statistically significant. Finally, the coefficient for graduate MPs is negative, but not statistically different from zero. Thus, after the reform of the electoral rules at the national level, rich and poor provinces continued to elect *a*-type MPs in the same proportion as they were doing before 1993, while rich jurisdictions began to elect *more p*-type MPs compared to poor ones, which is exactly the opposite behavior we observe at the municipal level. Again, this evidence of a different behavior of rich constituencies at national vs. municipal elections would be difficult to justify with the introduction of a new electoral law, reinforcing the idea that municipal fiscal autonomy plays a role in explaining our findings.

Controlling for covariates All the estimates described above have been obtained controlling only for municipal and year fixed effects. However, one might think that the observed change in the skills of mayors has been driven by factors varying across the years like the demand of services, the opportunity cost of becoming a politician, or the degree of political competition. To control for these stories, we add to model (7) a number of covariates which allow to control for time-varying municipal characteristics. In particular, we control for both the municipal population and the share of people older than 65 (to capture variables potentially affecting the demand for municipal services), the number of firms per capita at the provincial level (a proxy for the presence of entrepreneurs and the opportunity cost of becoming a politician), and the level of political competition, which we measure using a Herfindahl index on the shares of votes taken by all the Italian parties active at the municipal level³⁸. The results obtained after controlling for all these covariates are reported in columns (1), (3) and (5) of Table 5 for *a*-type politicians, and in columns (1) and (3) of Table 6 for *p*-type politicians³⁹. As shown, our results remain virtually unchanged in terms of sign, magnitude, and significance after the introduction of these additional covariates.

New Parties One could wonder that the effect on the selection of the political class might also be due to the new parties, Lega Nord and Forza Italia, that emerged after the political turmoil of the '90s, at the beginning of the so called "Second Republic". Perhaps, as these were new parties, they also selected completely new candidates, without past political experience, and this might explain our results. We do not control for the new political parties in the main specification in Tables 5-6, because these are potential dependent variables in our empirical model, and thus they should be treated as *bad controls*. However, as an exercise, we rerun all the regressions in Tables 5-6 adding dummy variables equal to one for those mayors affiliated to the new political parties that emerged during the years studied in our empirical analysis. As shown in Table A.4 (Appendix 1), the results described in Tables 5-6 remain entirely unchanged. The new parties did select different candidates, but their impact is not enough to explain the observed change in skills.

A differentiated shift in voters' preferences Another potential threat to the causal interpretation of our results is represented by the simultaneity between the municipal reforms and the political scandals at the beginning of the '90s. One might suspect that these might

³⁸Galasso and Nannicini (2011) stress the importance of competition for the quality of politicians in national elections in Italy. Notice that in order to avoid issues of endogeneity, we build the Herfindahl index using the shares of votes taken by political parties at the most recent European elections at the municipal level.

³⁹We do not report in the Tables the coefficients of these additional controls to save space. The complete set of results is available from the authors on request.

have affected voters differently in different areas of the country, leading only those living in the richest cities to ask for a change of the political ruling class⁴⁰. Our results would then be just capturing these differentiated shifts in voters' preferences fueled by the corruption scandal which wiped away the old ruling class.

First, the exercise on the national politicians discussed above suggests that this was not the case, as voters in the richer jurisdictions voted for *more p*-type MPs at the national elections rather than vice versa. However, as a second exercise, we consider here voting behavior at a national referendum on the funding of political parties, held on April 18, 1993 (e.g., Shugart and Wattenberg, 2001). In this referendum, Italians were asked if they wanted to abolish public funding for political parties⁴¹. Turnout was as high as 77%, close to the participation rate at national elections, and the proposal was approved in all Italian provinces, but with some variability across the country⁴². Thus, we can use the referendum outcome as a proxy for the willingness to change the political system in the different parts of the countries, and check whether a differentiated shift in preferences is responsible for our results. Specifically, we define a new variable measuring (at the provincial level⁴³) the share of voters in favour of abolishing public funding to parties (*%_referendum*), and interact this variable with *Decentr*⁴⁴. We insert this interaction in equation (9), where *Decentr* is interacted not only with our measure of fiscal autonomy, but also with a number of other variables potentially affecting the selections of local politicians (see below). Notice that the coefficient for the interaction related to referendum is never statistically significant at the usual confidence levels (Table 5, columns (2), (4) and (6); Table 6, columns (2) and (4)), while the coefficient for fiscal autonomy is still significant, maintains the expected sign, and strongly increases in magnitude. This suggests that a territorial shift in voters' preferences did not cause the observed change in the skills of mayors.

Social capital, supply of politicians, local politics Besides voters' preferences, one can still think of other alternative factors that might have changed their role after decentralization. As a general strategy, we discuss here these alternatives by adding new interaction terms with *Decentr* in model (9). First, one might think that the different endowment of social capital across Italian cities might still play a role in the selection of politicians, for instance via its influence on voters' preferences and electoral participation. We then consider the interaction between the number of non-profit organizations per capita (*#_no_profit*⁴⁵, a commonly used proxy for social capital, see Guiso et al., 2016) and *Decentr* in the augmented version of our baseline model. Second, while our model emphasizes the demand effects on local politicians, one might think that there are also supply-side effects driving our results. Perhaps, poor cities have

⁴⁰The reforms in the electoral rules were surely engineered by the popular outrage for the results of the "Clean Hands" trials, and these trials had their epicentre in Milano, one of the richest cities in our sample.

⁴¹The referenda were eight, all held simultaneously, ranging from partially allowing the use of drugs to reducing the role of some ministries such as Health, Agriculture and Tourism. We just pick the one more directly related to political parties.

⁴²The referendum on party funding, although approved, was never applied. The Parliament just substituted public financing with public refunds for electoral expenses, proxying these with a fixed sum of money for voter, so *de facto* reintroducing public funding to political parties.

⁴³There are no data for the results of referenda at the municipal level.

⁴⁴The variable *% referendum* enters the regression only via the interaction term because it is time invariant, and thus it is already captured by municipal fixed effects.

⁴⁵As for referendum, the variable *# no profit* enters the regression only via the interaction term because it is time invariant.

just less candidates of *a*-type available, and this explains why after the reform we observe more *p*-type mayors in these cities. We then consider two additional interactions with *Decentr*: the number of firms per capita at the provincial level to capture the availability of entrepreneurs, and the number of college graduated out of the adult population at the provincial level to capture the availability of professionals⁴⁶. Finally, the 1993 electoral reform affected many features of local politics, for instance giving birth to new parties and new coalitions, so that our explanation may miss these changes at the municipal level. We then consider an interaction between *Decentr* and the Herfindahl index of political competition at the local level discussed above⁴⁷.

Estimates for all these interactions are reported in Table 5, columns (2), (4) and (6), and Table 6, columns (2) and (4). As for civic capital, the coefficient for the interaction term is never statistically significant. The coefficients for both interactions capturing the availability of professionals and entrepreneurs are also almost always statistically insignificant, except for graduate in the *adm_index* case. More interestingly, coefficients for the interaction with political competition show a negative and statistically significant coefficient in two out of three models for *a*-type politicians, and are always positive and significant for *p*-type politicians. Hence, in line with findings by Galasso and Nannicini (2004) for national politicians, we confirm that less competition at the local level weakens the incentive to choose more useful mayors (i.e., *a*-type and less *p*-type) after the reform kicks in. Again, however, these alternative stories are unable to change our results.

A North-South divide? Southern cities are on average poorer than Northern cities, but Italian Southern regions are also well known to be characterized by a lower endowment of social capital and a larger presence of organized crime, differences that have been proved to impact on a wide range of economic and political outcomes (e.g., Putnam, 1993; Guiso et al., 2016) and that might also affect the selection of the local political class. Thus, one could suspect that our results might just capture the usual North-South divide, rather than being related to the strategic incentives for voters and candidates induced by the different wealth of the cities. Results on the number of non-profit organizations per capita as a proxy for civic capital suggest that this should not be the case. However, to provide a further check for this, we re-run (7) by first dropping all observations coming from the Southern regions⁴⁸ and then by dropping also all observations coming from the Centre⁴⁹, thus limiting our analysis only to municipalities in the North. Results are reported in Tables A.5 and A.6 (Appendix 1), both without (Panel A) and with control variables (Panel B). In spite of the reduced samples, all our results on the selection of mayors are confirmed in both cases and both in terms of sign and significance. This is further reassuring that our findings are not simply the result of the North-South divide.

Political campaign costs Political campaign costs might be an additional source of concern (e.g., Poutvaara and Takalo, 2007). As an alternative explanation for our results, one could argue that, following the electoral reform, the stronger emphasis on the role of mayors may

⁴⁶The variable % college also enters the regression only via the interaction term because it is time invariant. Notice that no other information on the type of graduates is available at municipal or provincial level.

⁴⁷We cannot define proper interaction terms between *Decentr* and the affiliation of mayors to new parties, in particular with Forza Italia, because this party did not exist before 1994. See the previous section for exercises controlling for the affiliation of mayors to the new parties.

⁴⁸The southern ordinary regions are Abruzzo, Basilicata, Calabria, Campania, Molise and Puglia.

⁴⁹The central regions are Marche, Lazio, Umbria and Toscana.

have led candidates to spend more money on their campaigns, and as richer people can spend more, this could explain why we observe more mayors coming from top administrative positions in the private sector in the richer cities. To be sure, this is not so much of a concern in Italy as in other countries, for different reasons. First, because political parties usually finance the electoral campaigns of their candidates, and political parties are in turn generously financed with public funds. Second, because municipal political campaigns in Italy are not very sophisticated, with candidates relying mostly on public rallies, direct mailing and TV appearances, all of which limit the budget required to run as a candidate for mayor (e.g., Kendall et al., 2015). Furthermore, in our specific case, administrative skills for mayors does not necessarily mean that candidates are rich: lawyers are excluded by our administrative professions, while school managers are included, and the latter are very likely to be on average poorer than the former. Nevertheless, while we cannot directly control for these effects because campaign expenditures at the municipal level are not publicly available, one could argue that the role of campaign costs is likely to be more important in the richest cities that are also *regional capitals*, such as Milano or Roma, rather than in smaller provincial cities such as Cremona or Viterbo belonging to the same two regions. As a control, we then re-run model (7) on a sample excluding all regional capitals where campaign costs might be more important. Results from this exercise are reported in Appendix 1, Table A.7. Our findings remain basically unchanged also in this case.

5.3 Additional robustness checks

Here we provide a number of additional further checks for our results, discussing more technical issues such as the choice of variables measuring fiscal autonomy and politicians' types, but also the definition of the sample. The general conclusion is that our main findings do not appear to be driven by our modelling choices.

Alternative measure of fiscal autonomy. Results so far have been obtained by relying on pre-reform GDP as a measure of fiscal autonomy. This solves the problem of endogeneity, but one might be worried by the fact that GDP does not exactly capture fiscal autonomy. Hence, as a robustness check, we consider here a more precise measure of fiscal autonomy defined as the ratio between municipal own revenues (in terms of taxes and fees) and total revenues (including grants from Central government and fees). We then average this ratio over the years 2000-2012 and use this average as a new definition of *FiscAutEstimates* of model (7) with this new measure of fiscal autonomy are reported in Table A.8 (Appendix 1) and are largely comparable in terms of sign and statistical significance with our baseline results. The choice of the fiscal autonomy measure is not driving our results.

Considering electoral years only. In the analysis so far we used all years, even non-electoral years. As we discussed above, this is because – before the change of the electoral rules – mayors often changed during the same electoral mandate between elections. However, one might be worried that this can artificially inflate the efficiency of our estimates, as an elected mayor may remain in charge for several years. To address this concern, in this section we drop non-electoral years and we re-estimate equation (7) by using only data for electoral years. As before the reform, mayors often changed, to build the dependent variable for the pre-reform years we use the average of the skills of all mayors in charge before the reform, weighted by the number

of years in office⁵⁰. Estimates are reported in Table A.9 (Appendix 1) and confirm in terms of sign, statistical significance, but also magnitude, our main findings. It is not an issue for our results whether we consider also non-electoral years.

The definition of Y^a and Y^p . Finally, one might be worried by the empirical definition of our indicators Y^a and Y^p . In our theoretical model, the a -type and the p -type are two mutually exclusive types of politicians, while in the empirical analysis so far we somewhat “mixed up” the two types, just considering one of the two relevant skills of the different politicians. Thus, for example, a mayor coming from a profession characterized by strong administrative skills would be considered an a -type even if she/he also had a long experience in politics. To check if this affects our results, we need a definition of Y^a and Y^p which is mutually exclusive. To this aim, we apply the following four mutually exclusive definition of Y^a and Y^p : 1) $Y^a[A] = 1$ if $adm_jobs[A] = 1$ and $years_pol_exp = 0$; 2) $Y^a[B] = 1$ if $adm_jobs[B] = 1$ and $years_pol_exp = 0$; 3) $Y^p[A] = 1$ if $adm_jobs[A] = 0$ and $years_pol_exp > 0$; 4) $Y^p[B] = 1$ if $adm_jobs[B] = 0$ and $years_pol_exp > 0$. We then re-run equation (7) by considering these new dependent variables. Results are reported in Table A.10 (Appendix 1) and they overwhelmingly confirm our baseline findings. Our evidence does not rely on the choice of modelling the two skills independently.

6 Voters’ welfare

6.1 Proxies for welfare

The main focus of this paper is on the skills of elected mayors. Our story is that administrative skills are more useful in rich communities, while political skills are more useful in poor ones. According to our theoretical model, this implies that a tax decentralization reform should impact on consumers’ welfare differently in rich and poor municipalities. Unfortunately, welfare is much more difficult to measure than skills. However, as we discussed above, having political skills means being better at attracting more funds, for instance manipulating the grant formula to the advantage of the municipality⁵¹; having administrative skills translates in improving the collection of local taxes or providing more efficiently local public services. Our model predicts that rich communities will elect candidates with administrative skills and welfare will improve; while electing candidates with political skills in poorer communities, even if rational for voters, will have uncertain effects on welfare. What we do here is then to provide evidence on the ex-post performance of mayors, looking for proxies of welfare related to administrative/managerial skills.

In particular, we proxy Y^w by selecting two measures, one on the revenue side and the other on the expenditure side, that are independent from the amount of resources that can be raised via the property tax and the PIT surcharge, hence avoiding the danger to attribute the better performance of a municipality to its higher tax resources. According to definition [A],

⁵⁰For example, if we have two mayors during the same electoral term, each covering half of the mandate, and for one the dummy variable $adm_jobs[A] = 1$, while for the other $adm_jobs[A] = 0$, the dependent variable used for this term is equal to 0.5.

⁵¹For instance, Carozzi and Repetto (2016) show that municipalities which are birthplaces of Italian members of Parliament receive more transfers from the central government. The result is driven by MPs who were born in a municipality outside their district of election. This reinforces the story about personal linkages between the incumbent mayor and MPs as a determinant of grants.

Y^w is the *speed of revenues collection* (*revenues_collection*), and closely captures the idea in Equation (2) above. This is defined as the ratio between municipal revenues that are *collected* and revenues that are *assessed* in a specific budget year. This ratio can be interpreted as an efficiency indicator, as more efficient municipal administrators should be able to cash a larger amount of the expected resources, and has been widely used in this sense in the literature (e.g., Gagliarducci and Paserman, 2012; Gagliarducci and Nannicini 2013). The higher the level of administrative skills, the higher the ratio. Data on *revenues_collection* are taken from Gagliarducci and Nannicini (2013) and are available for each year since 1993.

According to definition [B], Y^w is the percentage of *separate waste collection* (*%waste*)⁵², an indicator provided by Legambiente, an Italian independent environmental organization. Managing waste is an important task assigned to municipalities, on which mayors are easily evaluated by citizens and for which they are considered accountable⁵³. It is often used by international organizations (e.g., the United Nations Habitat Programme) as a measure of the quality of municipal governance. In Italy, waste management is not financed by tax revenues, but by a tariff to be paid for the service provided. This variable has the further advantage of being available for the Italian provincial capitals and each year since 1993, and to be a more reliable indicator of good governance for Italian municipalities than simple expenditure data. As for the speed of revenues collection, we expect that the higher the level of administrative skills, the higher the percentage of separate waste collection.

As we now consider a performance measure for elected mayors, we also account for the potential role of term limitation. In particular, we modify our model (7) by adding a dummy variable taking value one if the mayor is at his second term in office (and therefore is term limited, *TL*), and we interact it with our main variable $Decentr \times FiscAut$. This allows us to assess whether the effect of tax decentralization on the performance of mayors is due to re-election incentives or to a selection effect. In fact, while the behavior of a mayor during his first term in office (when he can be re-elected) can be driven by both his personal characteristics and by re-election incentives, in his second term in office his performance should be driven solely by his personal characteristics as re-election incentives are no longer in place.

6.2 Results

The results of this exercise are reported in Table 7: we consider the *speed of revenues collection* (*revenues_collection*) in column (1), while the dependent variable in column (2) is the *percentage of separate waste collection* (*%waste*). As shown in the Table, we do not find any statistically significant differential effect between rich and poor municipalities for mayors who can be re-elected (i.e., Term Limit=0). On the contrary, the tax decentralization reforms seem to have positively affected the performance of term limited mayors (i.e., mayors who cannot be re-elected, Term Limit = 1) in rich municipalities, compared to those elected in the poor ones. The estimated coefficients indicate that the ratio between collected and assessed revenues was around 11.5% points higher for second term mayors elected in rich municipalities. At the same time, the percentage of separate waste collection was 5.1% points higher for second term mayors elected in wealthy communities, compared to the mayors of poor municipalities.

⁵²Garbage can be collected uniformly or separately for the different sources of waste. Separate collection allows for recycling and it is therefore strongly supported by environmental organizations.

⁵³See, for instance, the echo in international media caused by the recent rubbish crisis in Naples in 2008, and in Palermo in 2012.

The results reported in Table 7 suggest that tax decentralization reforms have differentially affected performance across rich and poor municipalities through the selection of different types of politicians. When re-election incentives are in place (i.e., for first term mayors) we do not find differences in the effect of tax decentralization across rich and poor communities. Conversely, when re-election incentives are removed (as it is the case for second term mayors) and the only difference across municipalities is represented by the selection of different types of mayors, differences in the effect of tax decentralization between rich and poor communities emerge. This indicates that the positive effect of tax decentralization on the performance of mayors in rich municipalities, compared to the poor ones, works through the selection of more *a*-type and less *p*-type politicians, as suggested by our theoretical model.

7 Concluding remarks

The empirical literature on fiscal federalism supports the view that decentralization requires, in order to work properly, enough autonomous resources to finance local spending. The common explanation for this observed positive relationships between the degree of fiscal autonomy and decentralization outcomes is based on a “common pool” effect. We add to the literature by suggesting a complementary explanation based on a “selection effect” of local politicians: our theoretical model emphasizes that politicians endowed with administrative/managerial skills are more useful (hence, are more likely to be elected) in rich sub-national governments, while professional politicians are more useful in poor ones. We then provide supporting evidence for this argument in the case of the Italian decentralization reforms of the '90s involving municipalities. The empirical analysis provides evidence that after the 1993 reforms – when municipalities were made more fiscally autonomous and citizens were allowed to choose directly their main representative – in the richer cities voters increasingly elected mayors coming from top administrative professions and with less political experience, an effect that was later emphasized by the additional tax reform in 1999 (at unchanged electoral rules).

While this evidence is undisputable, attributing a causal meaning to this evidence is not straightforward, given all that happened in Italy at the beginning of the '90s. However, we provide additional support to our story reviewing a number of alternative potential explanations. Interestingly, for instance, the selection effect observed after the reforms influenced the specific skills of the local politicians, and not other personal characteristics or their general level of competence. The selection effect also clearly did not depend on a more general differentiated shift in political preferences, as we do not observe the same selection effect in rich areas for *national* politicians, and it also did not depend on other potential factors influencing local politics after the reform.

Given this additional support to our story, we are left with the view that the observed selection effect was due to the strategic incentives on voters and candidates provided by the different level of financial autonomy of the cities. Intuitively, in poorer cities, voters might simply not have enough incentives to choose a good administrator of local matters, as the basic task of the local politician is to ‘bribe’ the center in order to guarantee that central money keeps flowing to the community. Political skills and political connections are more important. In principle, of course, these incentives could be counteracted by rule-based formulas that eliminated all political discretion in the setting up of intergovernmental transfers. But the experience in several countries – including Italy – is that these rule-based transfers are very hard to implement in practice, and in any case they cannot be easily extended to all types of

transfers (for instance, investment projects and capital expenditure grants). As some political discretion is therefore unavoidable, in those countries and situations where local governments are heavily dependent on grants, it is then possible that decentralization will end up in better services only in rich communities. It would be interesting if future research could replicate our analysis to other countries interested by a decentralization process.

References

- [1] Ahmad E., Brosio G., eds. (2008). *Handbook of Fiscal Federalism*. Edward Elgar, London.
- [2] Alesina A., Paradisi M. (2017). "Political Budget Cycles: Evidence from Italian Cities". *Economics & Politics*
- [3] Ambrosanio M.F., Balduzzi P. and Bordignon M. (2015), "Economic Crisis and Fiscal Federalism in Italy" in Ahmad E., Bordignon M. and Brosio G. (eds), *Multi-level Finance and the Euro Crisis — causes and effects*, Edward Elgar.
- [4] Ashworth S. and Bueno de Mesquita E., (2008). "Electoral Selection, Strategic Challenger Entry, and the Incumbency Advantage." *The Journal of Politics*, Vol. 70, No. 4, October 2008, Pp. 1006–1025.
- [5] Atkinson, M. D., R. D., Enos, and S. J. Hill. (2009). "Candidate Faces and Election Outcomes: Is the Face–Vote Correlation Caused by Candidate Selection?" *Quarterly Journal of Political Science* 4: 229–49.
- [6] Angrist J. D., Pischke J.-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.
- [7] Autor, D. (2003). "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing". *Journal of Labor Economics*, 21(1): 1-42.
- [8] Baldini G., Legnante G. (2000). *Città al voto: i sindaci e elezioni comunali*. Il Mulino, Bologna.
- [9] Bardhan P., D. Mookherjee (2000). "Capture and Governance at Local and National Levels." *American Economic Review*, 90(2): 135-139.
- [10] Bardhan P., D. Mookherjee, eds. (2006) *Decentralization to Local Governments in Developing Countries: A Comparative Perspective*, MIT Press.
- [11] Beach B. and Jones D. B., (2016). "Business as usual: Politicians with business experience, government finances, and policy outcomes." *Journal of Economic Behavior & Organization*, Volume 131, Part A, November 2016, Pages 292-307.
- [12] Besley T. (2004). "Paying politicians: theory and evidence." *Journal of the European Economic Association* 2, 193-215.
- [13] Besley T. (2005). "Political Selection." *Journal of Economic Perspectives*, 19(3): 43–60.

- [14] Besley T. (2006). “Principled Agents? The Political Economy of Good Government”. Oxford University Press, Oxford.
- [15] Besley T., A. Case (1995). “Incumbent Behavior: Vote Seeking, Tax Setting and Yardstick Competition”, *American Economic Review*, 85(1): 25-45.
- [16] Besley, T. and Reynal-Querol M., (2011) “Do Democracies Select More Educated Leaders?,” *American Political Science Review*, 105 (2011), 552–556."
- [17] Besley T., Smart M. (2007). “Fiscal restraints and voter welfare”, *Journal of Public Economics*, 91(3-4): 755-773.
- [18] Bettin Lattes G., Magnier A. (1995). “I nuovi sindaci: come cambia una carriera politica”. *Rivista Italiana di Scienza Politica*, 25(1):91-118.
- [19] Bordignon M., Grembi V. and Piazza S. (2017). “Who do you Blame in Local Finance? An Analysis of Municipal Financing in Italy", *European Journal of Political Economy*.
- [20] Bordignon M., Cerniglia F., Revelli F. (2003). “In search of yardstick competition: a spatial analysis of Italian municipality property tax setting", *Journal of Urban Economics*, 54(2): 199-217.
- [21] Bordignon M., Colombo L., Galmarini U. (2008). “Fiscal federalism and lobbying". *Journal of Public Economics*, 92(12): 2288-2301.
- [22] Bordignon M., Nannicini T., Tabellini G. (2016). “Moderating political extremism", *American Economic Review*, 106 (8), 2349-70.
- [23] Bordignon M., Turati G. (2009). “Bailing out expectations and public health expenditure", *Journal of Health Economics*, 28 (2), 305-21.
- [24] Brollo F., and Nannicini T., (2012). “Tying Your Enemy’s Hands in Close Races: The Politics of Federal Transfers in Brazil." *American Political Science Review*, Vol. 106, No. 4.
- [25] Brollo F., Nannicini T., Perotti R., Tabellini G. (2013). “The Political Resource Curse". *American Economic Review*, 103 (2013), 1759-1796.
- [26] Cai H. and Treisman D. (2005). “Does Competition for Capital Discipline Governments? Decentralization, Globalization and Public Policy.” *American Economic Review*, 95(3): 817-830.
- [27] Cameron A. C., Gelbach J. B., Miller D. L. (2011). “Robust Inference With Multiway Clustering”. *Journal of Business and Economic Statistics*, 29(2): 238-249.
- [28] Card D. (1992). “Using Regional Variation in Wages to Measure the Effects of the Federal Minimum Wage”. *Industrial and Labor Relations Review*, 46(1): 22-37.
- [29] Carozzi F. and Repetto L. (2016). “Sending the pork home: birth town bias in transfers to Italian municipalities ”. *Journal of Public Economics*, 134(2): 42-52.

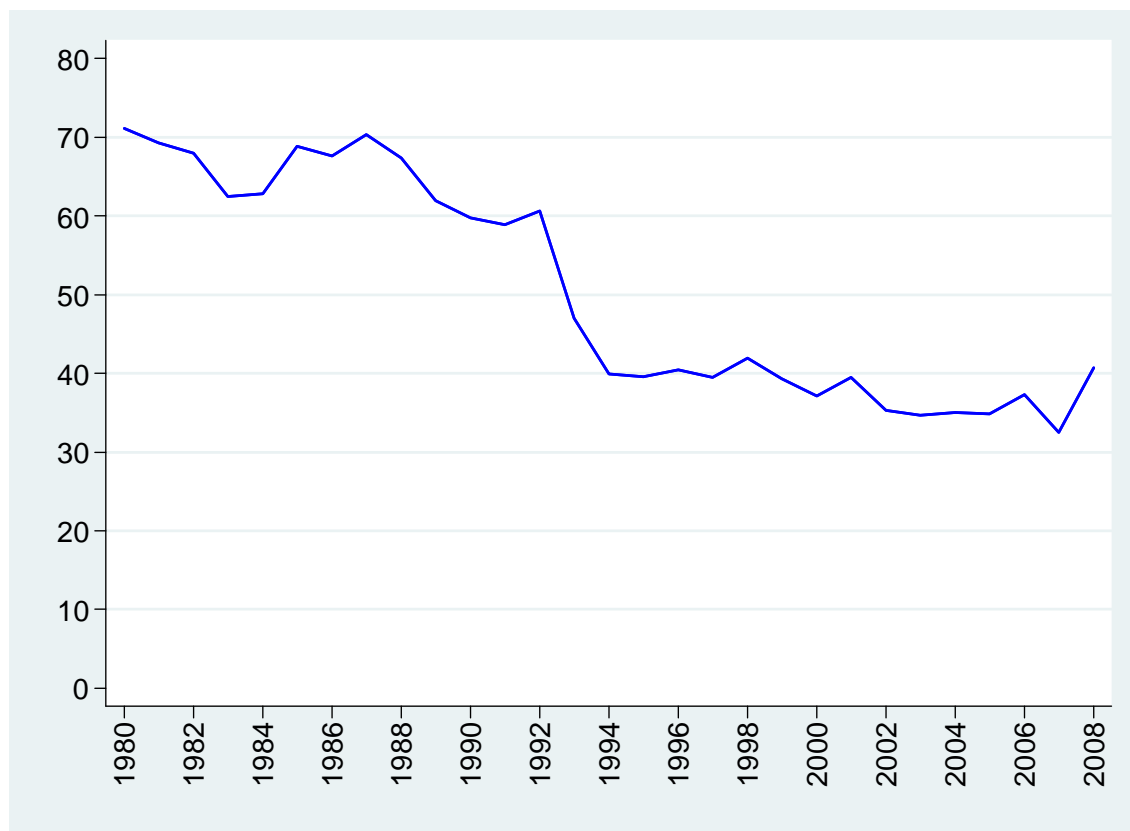
- [30] Caselli F. and Morelli M. (2004). "Bad politicians". *Journal of Public Economics*, 88, 759-782.
- [31] Dal Bo, Dal Bo, and Di Tella, (2006). "Plata o Plomo? Bribe and Punishment in a Theory of Political Influence." *American Political Science Review*, 100 (1): 41–53.
- [32] Dal Bo' E., Finan F., Folke O. and Rickne J. (2017). "Who becomes a politician?" *Quarterly Journal of Economics*.
- [33] Daniele G. (2017). "Strike one to educate one hundred: Organized crime, political selection and politicians' ability." *Journal of Economic Behavior & Organization* (2017).
- [34] Daniele G. and Geys B. (2015). "Organised crime, institutions and political quality: empirical evidence from Italian municipalities." *Economic Journal*, 125: F233–F255.
- [35] Dahlberg M., E. Mörk, J. Rattsø and H. Ågren (2008). "Using a discontinuous grant rule to identify the effect of grants on local taxes and spending". *Journal of Public Economics*, 92(12), 2320-2335.
- [36] D'Alimonte R. (2001). "Mixed electoral rules, partisan realignment, and party system change in Italy", in Shugart M. S., M. P. Wattenberg (eds.), *Mixed Member Electoral Systems: The Best of Both Worlds?*, Oxford University Press.
- [37] Eyraud L., Lusinyan L. (2013). "Vertical fiscal imbalances and fiscal performance in advanced economies". *Journal of Monetary Economics*, 60(5): 571-587.
- [38] Ferraz C., Finan F. (2009). "Motivating Politicians: the impacts of monetary incentives on quality and performance." NBER Working Paper No. 14906.
- [39] Fisman, R. and R. Gatti (2002), "Decentralization and Corruption: Evidence from US Federal Transfer Programs", *Public Choice*, 113: 25-35.
- [40] Folke, O., and J. Rickne (2016) "The Glass Ceiling in Politics Formalization and Empirical Tests," *Comparative Political Studies*, , 567–599.
- [41] Francese, M., Piacenza, M., Romanelli, M., Turati, G., (2014) "Understanding inappropriateness in health spending: The role of regional policies and institutions in caesarean deliveries," *Regional Science and Urban Economics*, 49(C), 262–277.
- [42] Gagliarducci S., Nannicini T. (2013). "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection". *Journal of the European Economic Association*, 11: 369-398.
- [43] Gagliarducci S., Nannicini T. and P. Naticchioni (2010). "Moonlighting Politicians" *Journal of Public Economics*, 94(9-10), 688-699.
- [44] Gagliarducci S., Paserman D. (2012). "Gender Interactions within Hierarchies: Evidence from the Political Arena" *Review of Economic Studies*, 79(3), 1021-1052.
- [45] Galasso V., Nannicini T. (2011). "Competing on good politicians." *American Political Science Review*, 105(1): 79-99.

- [46] Galiani, S., P. Gertler and E. Schargrodsy (2008). "School Decentralization: Helping the Good Get Better, but Leaving the Poor Behind." *Journal of Public Economics*, 92(10-11): 2106 -2120.
- [47] Gehlbach S., Sonin K. and Zhuravskaya E., (2010). "Businessman candidates." *American Journal of Political Science*, Volume 54, Issue 3, Pages 718-736, July 2010.
- [48] Guiso L., Sapienza P. and Zingales L. (2011). "Civic Capital as the Missing Link". In J. Benhabib, A. Bisin and M. Jackson (eds.), *Handbook of Social Economics*, Vol. 1, 417-480, Elsevier.
- [49] Guiso L., Sapienza P. and Zingales L. (2016). "Long-term Persistence." *Journal of the European Economic Association*.
- [50] Hayeon Joo H. and Lee J., (2018). "Encountering female politicians." *Journal of Economic Behavior & Organization*, Volume 151, July 2018, Pages 88-122.
- [51] Hines J. R. and R. H. Thaler (1995). "The Flypaper Effect". *Journal of Economic Perspectives*, 9(4): 217-226.
- [52] Imbens G. W. and J. M. Wooldridge (2009). "Recent developments in the econometrics of program evaluation". *Journal of Economic Literature*, 47(1): 5-86.
- [53] Katz R. S. (2001). "Reforming the Italian electoral law, 1993", in Shugart M. S., M. P. Wattenberg (eds.), *Mixed Member Electoral Systems: The Best of Both Worlds?*, Oxford University Press.
- [54] Kendall C., Nannicini T., Trebbi F. (2015). "How Do Voters Respond to Information? Evidence from a Randomized Campaign". *American Economic Review*, 105(1): 322-353.
- [55] Key, V. O., (1956). "American State Politics: An Introduction." New York: Alfred A. Knopf.
- [56] Larcinese, V., L. Rizzo, and C. Testa, (2006). "Allocating the U.S.Federal Budget to the States: The Impact of the President." *Journal of Politics* 68: 447-56.
- [57] Levitt, D.S., and J.M. Snyder. (1995). "Political Parties and the Distribution of Federal Outlays." *American Journal of Political Science*, 39: 958-80.
- [58] Lockwood B. (2015). "Fiscal Decentralization: A Political Economy Perspective", in E. Ahmad and G. Brosio (eds.), *The Handbook of Fiscal Federalism*, Edward Elgar.
- [59] Mattozzi A. and Merlo A. (2008). "Political careers or career politicians?" *Journal of Public Economics*, 92(3-4): 597-608.
- [60] Myerson, Roger B., (2006). "Federalism and incentives for success of democracy." *Quarterly Journal of Political Science* 1, 3-23.
- [61] Pasquino G. (2006). *I sistemi elettorali. Il Mulino*, Bologna.
- [62] Persson T., Tabellini G. (1994). "Does centralization increase the size of government?" *European Economic Review*, 38(3-4): 765-773.

- [63] Persson T., Tabellini G. (2000). "Political Economics: Explaining Economic Policy". Cambridge, MA: MIT Press.
- [64] Poutvaara P., Takalo T. (2007). "Candidate quality". *International Tax and Public Finance*, 14: 7-27.
- [65] Putnam R. (1993) Making Democracy Work. Civic tradition in Modern Italy. Princenton University Press.
- [66] Reinikka R. and J. Svensson (2004). "Local Capture: Evidence from a Central Government Transfer Program in Uganda." *Quarterly Journal of Economics*, 119(2): 679-705.
- [67] Rodden J. (2002). "The dilemma of fiscal federalism: grants and fiscal performance around the world." *American Journal of Political Science*, vol. 46, No.3, July 2002
- [68] Rodden J. (2006). "The Political Economy of Federalism" in Weingast B. and D. Wittman (eds.), Oxford Handbook of Political Economy, Oxford University Press.
- [69] Rodden, J. A., Eskeland G. S., and Jennie L. (2003). "Fiscal Decentralization and the Challenge of Hard Budget Constraints." Cambridge, MA: MIT Press.
- [70] Rogoff K., Siebert A. (1988). "Elections and Macroeconomic Policy Cycles". *Review of Economic Studies*, 55:1-16.
- [71] Shugart, M. S., M. E. Valdini, and K. Suominen, (2005). "Looking for Locals: Voter Information Demands and Personal Vote-earning Attributes of Legislators under Proportional Representation." *American Journal of Political Science* 49 (2): 437-49.
- [72] Shugart M. S., M. P. Wattenberg (eds.) (2001). "Mixed Member Electoral Systems: The Best of Both Worlds?", Oxford: Oxford University Press.
- [73] Treisman D. (2007). "The Architecture of Government. Rethinking Political Decentralization", Cambridge: Cambridge University Press.
- [74] Weingast, B. R. (2009). "Second generation fiscal federalism: The implications of fiscal incentives" *Journal of Urban Economics* 65 (2009) 279-293.
- [75] Weingast, B.R., Sheple K.A. and Johnsen, C. (1981) "The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics." *Journal of Political Economy*, 89 (4):642-64.

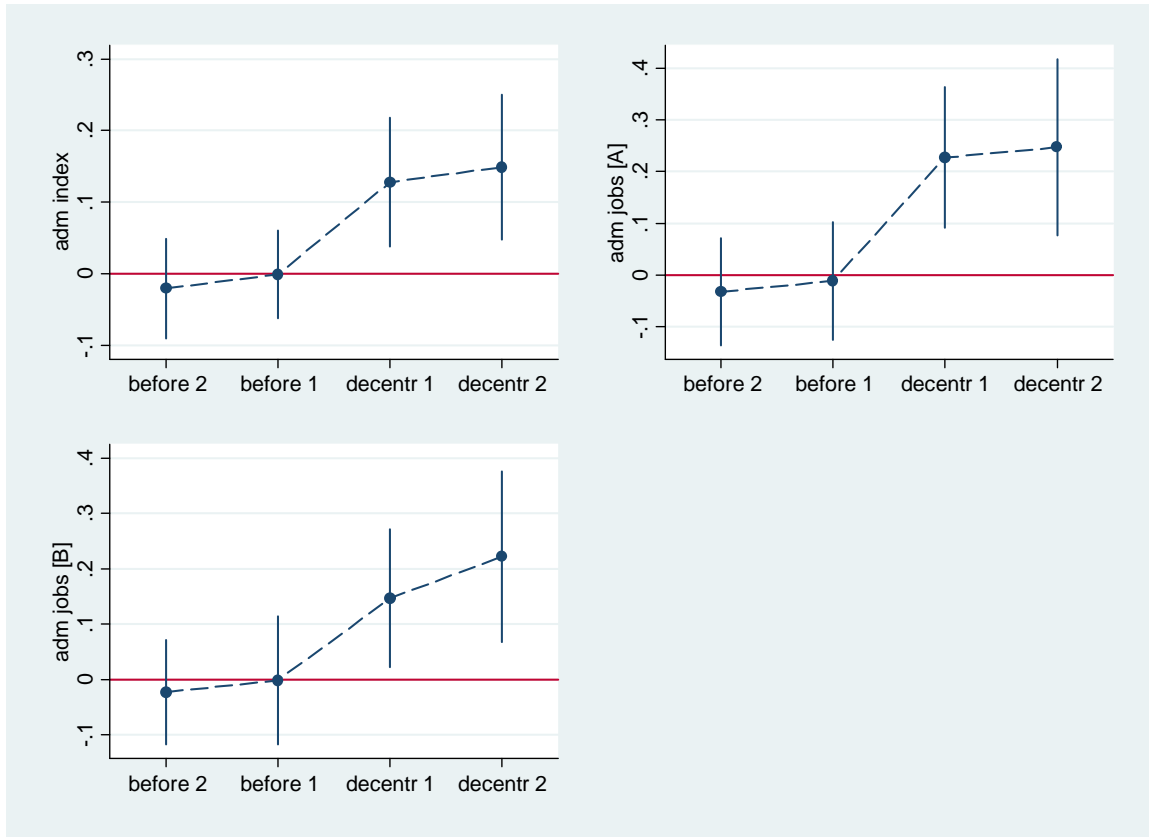
8 Tables and Figures

Figure 1: Central government transfers to Municipalities % current revenues (1980-2008)



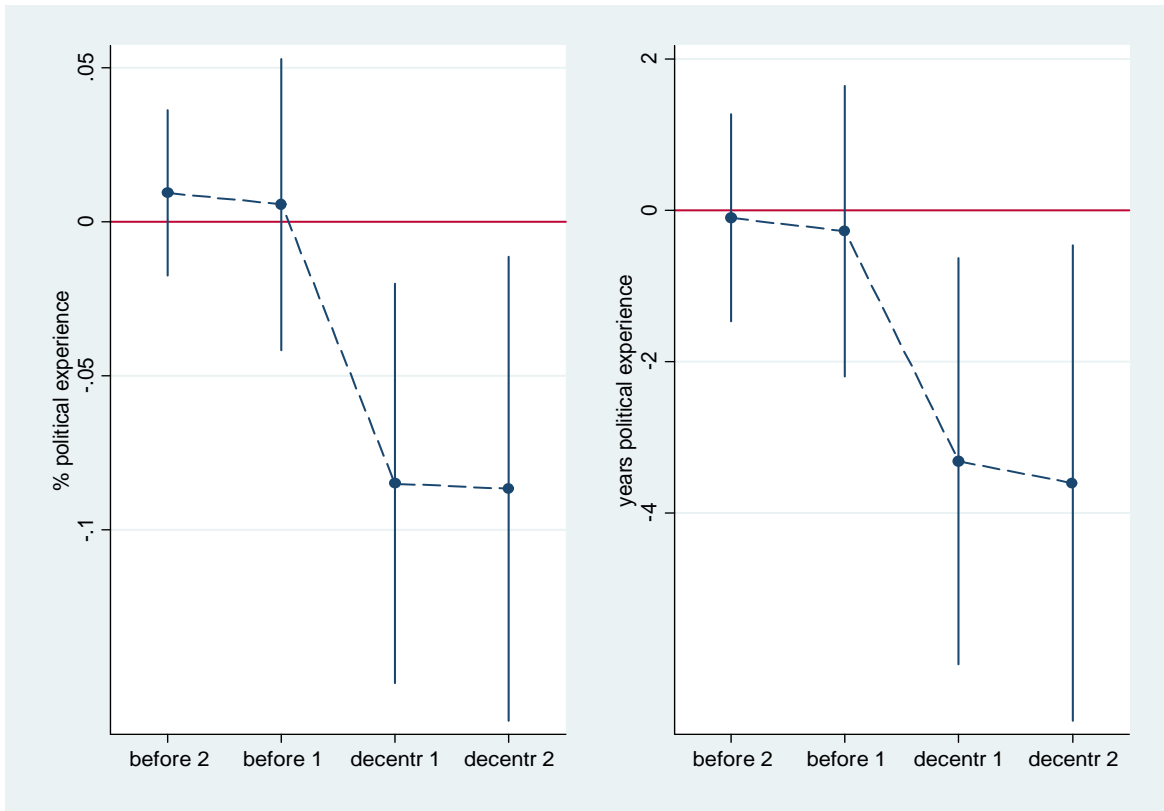
Source: ISTAT, Conto economico delle amministrazioni comunali.

Figure 2: Effect of decentralization on administrative skills



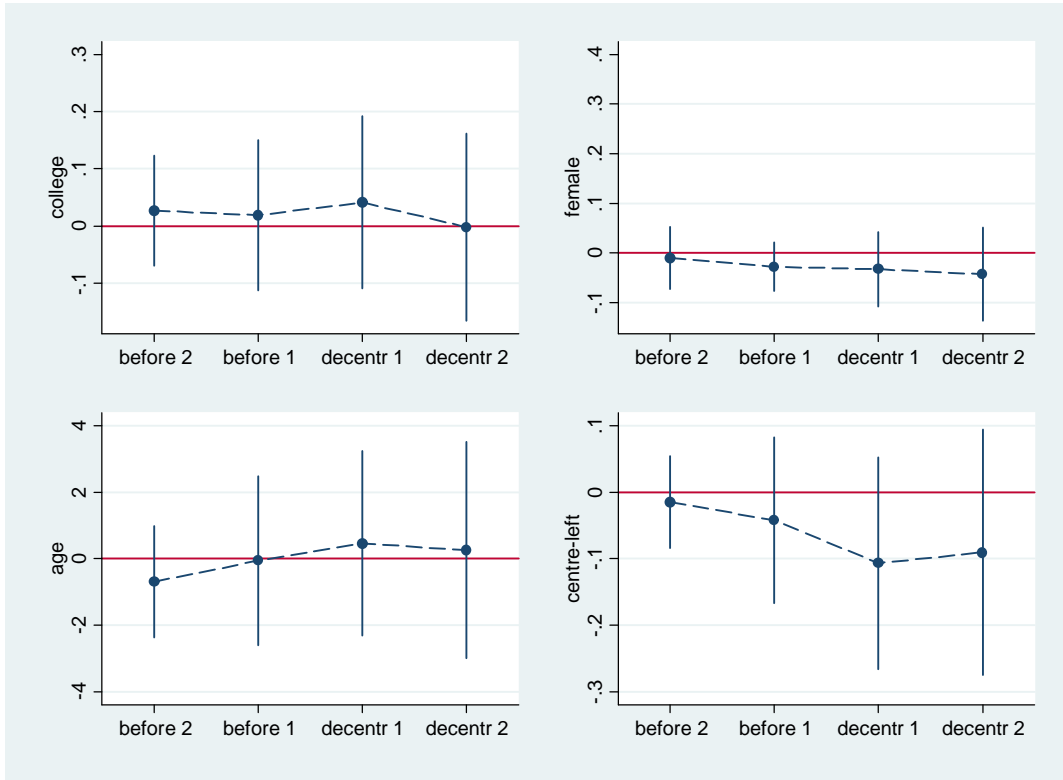
Variable definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants.

Figure 3: Effect of decentralization on political skills



Variable definition: % political experience: percentage of working years a candidate had already dedicated to politics before becoming mayor; years political experience: years of past political experience before becoming mayor.

Figure 4: Effect of decentralization on other characteristics



Variable definition: college=1 for graduate mayors; female=1 for female mayor; age: age of the mayor; centre-left=1 for mayor from a centre-left coalition.

Table 1: Impact of decentralization on types of politicians (estimates of δ)
Baseline model

	(1)	(2)	(3)
Panel A: Administrative Skills			
	adm_index	adm_jobs [A]	adm_jobs [B]
decentr \times FiscAut	0.137*** (0.039)	0.246*** (0.052)	0.163*** (0.049)
obs.	1,304	1,304	1,304
Panel B: Political Skills			
	% pol_exp	years_pol_exp	
decentr \times FiscAut	-0.091*** (0.022)	-3.168*** (0.856)	
obs.	1,330	1,330	
Year FE	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes
Covariates	No	No	No

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: Panel A: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants. Panel B: % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993.

Table 2: Impact of decentralization on types of politicians (estimates of δ_1 and δ_2)
Two decentralization reforms, administrative skills

	(1)	(2)	(3)	(4)	(5)	(6)
Administrative Skills						
	adm_index	adm_index	adm_jobs [A]	adm_jobs [A]	adm_jobs [B]	adm_jobs [B]
decentr 1 \times FiscAut	0.135*** (0.039)	0.128*** (0.046)	0.244*** (0.052)	0.227*** (0.069)	0.155*** (0.049)	0.147** (0.063)
decentr 2 \times FiscAut	0.155*** (0.045)	0.149*** (0.052)	0.262*** (0.072)	0.247*** (0.087)	0.230*** (0.065)	0.222*** (0.079)
1 year before \times FiscAut		-0.001 (0.031)		-0.012 (0.058)		-0.001 (0.059)
2 years before \times FiscAut		-0.021 (0.035)		-0.032 (0.053)		-0.023 (0.048)
obs.	1,304	1,304	1,304	1,304	1,304	1,304
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	No	No	No	No	No

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants. Independent variables definition: decentr 1: dummy variable equal to 1 after the introduction of the property tax (ICI); decentr 2: dummy variable equal to 1 after the introduction of the income surcharge tax (IRPEF); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993; 1 year before: dummy variable equal to 1 one year before the introduction of the property tax (ICI); 2 years before: dummy variable equal to 1 two years before the introduction of the property tax (ICI).

Table 3: Impact of decentralization on types of politicians (estimates of δ_1 and δ_2)
Two decentralization reforms, political skills

	(1)	(2)	(3)	(4)
Political Skills				
	% pol_exp	% pol_exp	years_pol_exp	years_pol_exp
decentr 1 \times FiscAut	-0.091*** (0.022)	-0.085** (0.033)	-3.133*** (0.841)	-3.318** (1.369)
decentr 2 \times FiscAut	-0.092*** (0.029)	-0.087** (0.038)	-3.436*** (1.112)	-3.605** (1.602)
1 year before \times FiscAut		0.006 (0.024)		-0.278 (0.980)
2 years before \times FiscAut		0.009 (0.014)		-0.102 (0.698)
obs.	1,330	1,330	1,330	1,330
Year FE	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes
Covariates	No	No	No	No

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr 1: dummy variable equal to 1 after the introduction of the property tax (ICI); decentr 2: dummy variable equal to 1 after the introduction of the income surcharge tax (IRPEF); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993; 1 year before: dummy variable equal to 1 one year before the introduction of the property tax (ICI); 2 years before: dummy variable equal to 1 two years before the introduction of the property tax (ICI).

Table 4: Impact of decentralization on other mayors' characteristics (estimates of δ)
 Placebo test

	(1)	(2)	(3)	(4)
Other mayors' characteristics				
	college	female	age	centre-left
decentr \times FiscAut	0.018	-0.016	0.677	-0.077
	(0.051)	(0.029)	(0.927)	(0.065)
obs.	1,330	1,330	1,330	1,335
Year FE	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes
Covariates	No	No	No	No

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: college: dummy variable equal to 1 if the mayor has a college degree; female: dummy variable equal to 1 if the mayor is a woman; age: age of the mayor; centre-left: dummy variable equal to 1 if the mayor comes from the centre-left coalition. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993.

Table 5: Impact of decentralization on types of politicians (estimates of δ)
Adding control variables and interaction terms

	(1)	(2)	(3)	(4)	(5)	(6)
Administrative Skills						
	adm_index	adm_index	adm_jobs [A]	adm_jobs [A]	adm_jobs [B]	adm_jobs [B]
decentr \times FiscAut	0.136*** (0.037)	0.288*** (0.107)	0.243*** (0.052)	0.570*** (0.173)	0.167*** (0.049)	0.323* (0.182)
decentr \times # firms		-0.026 (0.031)		-0.023 (0.045)		0.018 (0.047)
decentr \times pol. comp.		-0.000** (0.000)		-0.000* (0.000)		-0.000 (0.000)
decentr \times % college		-0.017** (0.007)		-0.008 (0.014)		-0.011 (0.016)
decentr \times # no profit		-0.001 (0.003)		-0.002 (0.007)		0.001 (0.006)
decentr \times % referendum		0.005 (0.003)		0.009 (0.005)		0.004 (0.005)
obs.	1,304	1,304	1,304	1,304	1,304	1,304
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993; # firms: number of firms per capita at the provincial level; pol. comp.: herfindal index of political competition; % college: percentage of municipal adult population with a college degree; # no profit: number of no profit associations at the municipal level; % referendum: % of yes at referendum on party funding. Additional control variables added to the regression but not reported in the Table: in all columns: municipal population; share of municipal population older than 65 years old; # firms not interacted with decentr, pol. comp. not interacted with decentr. In columns (2), (4) and (6): municipal population interacted with decentr; share of municipal population older than 65 years old interacted with decentr. The variables % college, % referendum and # no profit enter only in interaction terms because they are time invariant, so they are already captured by municipal fixed effects.

Table 6: Impact of decentralization on types of politicians (estimates of δ)
Adding control variables and interaction terms

	(1)	(2)	(3)	(4)
	Political Skills			
	% pol_exp	% pol_exp	years_pol_exp	years_pol_exp
decentr \times FiscAut	-0.092*** (0.021)	-0.243** (0.096)	-3.093*** (0.831)	-6.446* (3.886)
decentr \times # firms		-0.020 (0.022)		-0.972 (0.869)
decentr \times pol. comp.		0.000** (0.000)		0.008** (0.003)
decentr \times % college		-0.005 (0.006)		0.031 (0.231)
decentr \times # no profit		0.003 (0.004)		0.001 (0.122)
decentr \times % referendum		-0.003 (0.002)		-0.115 (0.087)
obs.	1,304	1,304	1,304	1,304
Year FE	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993; # firms: number of firms per capita at the provincial level; pol. comp.: herfindal index of political competition; % college: percentage of municipal adult population with a college degree; # no profit: number of no profit associations at the municipal level; % referendum: % of yes at referendum on party funding. Additional control variables added to the regression but not reported in the Table: in all columns: municipal population; share of municipal population older than 65 years old; # firms not interacted with decentr, pol. comp. not interacted with decentr. In columns (2), (4) and (6): municipal population interacted with decentr; share of municipal population older than 65 years old interacted with decentr. The variables % college, % referendum and # no profit enter only in interaction terms because they are time invariant, so they are already captured by municipal fixed effects.

Table 7: Impact of decentralization on mayors' performance

	(1)	(2)
Mayors' performance		
	revenues collection	waste collection
decentr \times FiscAut, Term Limit=0	0.566 (1.044)	0.547 (0.697)
decentr \times FiscAut, Term Limit=1	11.491** (4.715)	5.124** (2.191)
Difference	10.925** (4.754)	4.577** (1.976)
obs.	817	796
Year FE	Yes	Yes
Municipal FE	Yes	Yes
Covariates	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: revenues collection is the ratio between collected and assessed revenues; waste collection: percentage of separate waste collection. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993. Additional control variables added to the regression but not reported in the Table: Term Limit: dummy variable equal to 1 if the mayor cannot be re-elected (i.e. second term mayor); municipal population; share of municipal population older than 65 years old; # firms: number of firms per capita at the provincial level ; pol. comp.: herfindal index of political competition. All the variables have been interacted with Term Limit.

9 Appendix 1 [For Online Publication]

Table A.1: Descriptive statistics (1988-2002)

Variable	Obs	Mean	Std. Dev.	Min	Max	Source
adm_index	1304	3.898	0.286	3.296	4.327	Italian Ministry of Domestic Affairs,
adm_jobs[A]	1304	0.292	0.455	0	1	"Anagrafe degli amministratori locali"
adm_jobs[B]	1304	0.212	0.409	0	1	
% pol_exp	1330	0.263	0.211	0	0.93	Italian Ministry of Domestic Affairs,
years_pol_exp	1330	8.529	7.545	0	56	"Anagrafe degli amministratori locali"; OPENOPOLIS; historical archive "Corriere della Sera" and "La Repubblica"; Italian and European Parliament
college	1330	0.675	0.468	0	1	Italian Ministry of Domestic Affairs,
female	1330	0.053	0.225	0	1	"Anagrafe degli amministratori locali"
age	1330	50.648	9.312	29	85	
centre-left	1335	0.53	0.49	0	1	Italian Ministry of Domestic Affairs,
waste collection	796	11.082	9.856	0	52.1	"Anagrafe degli amministratori locali"; OPENOPOLIS; historical archive "Corriere della Sera" and "La Repubblica"
revenues collection	817	67.83	13.53	13.6	91.4	Legambiente
GDPpc	1335	16298	3731	8851	26041	Gagliarducci and Nannicini (2013)
population	1335	171622	338180	15008	2800000	Istituto Tagliacarne
% over 65	1335	17.994	4.022	7.690	27.88	ISTAT
# no profit	1335	0.006	0.002	0.002	0.012	
# firms	1335	7.741	1.787	4.050	13.149	Italian Chambers of Commerce
pol. comp.	1335	1963.897	439.553	1045.534	3657.212	Italian Ministry of Domestic Affairs,
% referendum	1335	90.1	3.2	81.6	96.6	"archivio storico delle elezioni"; ITANES
term limit	1335	0.155	0.362	0	1	Italian Ministry of Domestic Affairs,
						"Anagrafe degli amministratori locali"

Table A.2: Placebo tests (1988-1992)

	(1)	(2)	(3)
Panel A: Administrative Skills			
	adm_index	adm_jobs [A]	adm_jobs [B]
fake_decentr \times FiscAut	-0.038 (0.069)	-0.117 (0.106)	-0.025 (0.096)
obs.	422	422	422
Panel B: Political Skills			
	% pol_exp	years_pol_exp	
fake_decentr \times FiscAut	-0.015 (0.053)	0.699 (1.662)	
obs.	442	442	
Year FE	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes
Covariates	No	No	No

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: Panel A: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants. Panel B: % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: fake_decentr: dummy variable equal to 1 starting from 1990; FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993.

Table A.3: Characteristics of Members of Parliament

Panel A: Administrative Skills	
adm_index	0.070 (0.046)
adm_jobs [A]	0.043 (0.069)
adm_jobs [B]	0.050 (0.059)
obs.	413
Panel B: Political Skills	
politician	0.067 (0.042)
# past terms	0.801*** (0.224)
# past years	3.085*** (0.839)
other offices	-0.031 (0.066)
obs.	415
Panel C: Graduate	
graduate	-0.054 (0.064)
obs.	415

Notes. Standard errors clustered at provincial level. Sig. Lev.: *** 1%, ** 5%, * 10%. Provincial and year of election FE included. Dependent variables definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; politician=1 for professional politicians and trade unionists; # past terms: number of past terms as member of parliament; # past years: number of past years as member of parliament; other offices=1 for MPs elected at other levels of government in the past; graduate=1 for graduate MPs.

Table A.4: Impact of decentralization on types of politicians (estimates of δ)
Controlling for new political parties

	(1)	(2)	(3)	(4)	(5)
	adm_index	adm_jobs [A]	adm_jobs [B]	% pol_exp	years_pol_exp
decentr \times FiscAut	0.128*** (0.038)	0.224*** (0.054)	0.161*** (0.049)	-0.083*** (0.019)	-2.812*** (0.759)
Forza Italia	0.162*** (0.046)	0.249** (0.110)	0.132 (0.118)	-0.118*** (0.040)	-3.308*** (1.268)
Lega Nord	0.105** (0.053)	0.277* (0.144)	0.081 (0.136)	-0.172*** (0.048)	-6.498*** (1.621)
obs.	1304	1304	1304	1330	1330
Year FE	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993; Forza Italia: =1 if mayor affiliated to Forza Italia; Lega Nord: =1 if mayor affiliated to Lega Nord. Additional control variables added to the regression but not reported in the Table: municipal population; share of municipal population older than 65 years old; # firms not interacted with decentr, political competition. The variables % college, % referendum and # no profit are time invariant, so they are already captured by municipal fixed effects.

Table A.5: Impact of decentralization on types of politicians (estimates of δ)
Municipalities in centre-north regions only

	(1)	(2)	(3)	(4)	(5)
	adm_index	adm_jobs [A]	adm_jobs [B]	% pol_exp	years_pol_exp
Panel A: without control variables					
decentr \times FiscAut	0.136***	0.269***	0.170***	-0.096***	-2.768***
	(0.042)	(0.063)	(0.061)	(0.022)	(0.732)
obs.	929	929	929	944	944
Panel B: with control variables					
decentr \times FiscAut	0.143***	0.268***	0.173***	-0.099***	-2.770***
	(0.039)	(0.059)	(0.059)	(0.021)	(0.695)
obs.	929	929	929	944	944
Year FE	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993. Additional control variables added to the regression but not reported in the Table (Panel B only): municipal population; share of municipal population older than 65 years old; # firms not interacted with decentr, political competition. The variables % college, % referendum and # no profit are time invariant, so they are already captured by municipal fixed effects. Municipalities in centre-north regions only (i.e. Emilia Romagna, Liguria, Lombardia, Piemonte, Veneto, Lazio, Marche, Toscana, Umbria).

Table A.6: Impact of decentralization on types of politicians (estimates of δ)
Municipalities in north regions only

	(1)	(2)	(3)	(4)	(5)
	adm_index	adm_jobs [A]	adm_jobs [B]	% pol_exp	years_pol_exp
Panel A: without control variables					
decentr \times FiscAut	0.182***	0.311***	0.209***	-0.095***	-2.467***
	(0.047)	(0.077)	(0.075)	(0.026)	(0.851)
obs.	594	594	594	600	600
Panel B: with control variables					
decentr \times FiscAut	0.176***	0.287***	0.199***	-0.086***	-2.355**
	(0.044)	(0.068)	(0.069)	(0.027)	(0.964)
obs.	594	594	594	600	600
Year FE	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993. Additional control variables added to the regression but not reported in the Table (Panel B only): municipal population; share of municipal population older than 65 years old; # firms not interacted with decentr, political competition. The variables % college, % referendum and # no profit are time invariant, so they are already captured by municipal fixed effects. Municipalities in north regions only (i.e. Emilia Romagna, Liguria, Lombardia, Piemonte, Veneto).

Table A.7: Impact of decentralization on types of politicians (estimates of δ)
Role of electoral campaigns (excluding regional capitals)

	(1)	(2)	(3)	(4)	(5)
	adm_index	adm_jobs [A]	adm_jobs [B]	% pol_exp	years_pol_exp
Panel A: without control variables					
decentr \times FiscAut	0.147*** (0.034)	0.261*** (0.049)	0.261*** (0.049)	-0.090*** (0.022)	-2.840*** (0.910)
obs.	1,081	1,081	1,081	1,105	1,105
Panel B: with control variables					
decentr \times FiscAut	0.142*** (0.034)	0.253*** (0.049)	0.253*** (0.049)	-0.089*** (0.023)	-2.724*** (0.908)
obs.	1,081	1,081	1,081	1,105	1,105
Year FE	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993. Additional control variables added to the regression but not reported in the Table (Panel B only): municipal population; share of municipal population older than 65 years old; # firms not interacted with decentr, political competition. The variables % college, % referendum and # no profit are time invariant, so they are already captured by municipal fixed effects. Regional capitals excluded from the sample.

Table A.8: Impact of decentralization on types of politicians (estimates of δ)
Alternative measure for fiscal autonomy

	(1)	(2)	(3)		
	adm_index	adm_jobs [A]	adm_jobs [B]	% pol_exp	years_pol_exp
Panel A: without control variables					
decentr \times FiscAut	0.319*** (0.086)	0.604*** (0.110)	0.424*** (0.110)	-0.210*** (0.046)	-6.955*** (1.613)
obs.	1,304	1,304	1,304	1,330	1,330
Panel B: with control variables					
decentr \times FiscAut	0.305*** (0.085)	0.586*** (0.106)	0.416*** (0.109)	-0.209*** (0.047)	-6.894*** (1.608)
obs.	1,304	1,304	1,304	1,330	1,330
Year FE	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to the average level of financial autonomy of the municipality for the years 2000-2012. Financial autonomy is equal to the ratio between municipal own revenues (taxes+fees) divided by total municipal revenues (taxes+fees+grants from higher levels of government). Additional control variables added to the regression but not reported in the Table (Panel B only): municipal population; share of municipal population older than 65 years old; # firms not interacted with decentr, political competition. The variables % college, % referendum and # no profit are time invariant, so they are already captured by municipal fixed effects. Regional capitals excluded from the sample.

Table A.9: Impact of decentralization on types of politicians (estimates of δ)

Term level data					
	(1)	(2)	(3)	(4)	(5)
	adm_index	adm_jobs [A]	adm_jobs [B]	% pol_exp	years_pol_exp
Panel A: without control variables					
decentr \times FiscAut	0.207***	0.276***	0.196**	-0.080*	-3.409**
	(0.061)	(0.104)	(0.096)	(0.041)	(1.640)
obs.	414	414	414	416	416
Panel B: with control variables					
decentr \times FiscAut	0.212***	0.270**	0.223**	-0.090**	-3.132*
	(0.056)	(0.113)	(0.089)	(0.043)	(1.736)
obs.	414	414	414	416	416
Term FE	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: adm_index: log of the index for administrative skills (ISFOL); adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % pol_exp = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993. Additional control variables added to the regression but not reported in the Table (Panel B only): municipal population; share of municipal population older than 65 years old; # firms not interacted with decentr, political competition. The variables % college, % referendum and # no profit are time invariant, so they are already captured by municipal fixed effects.

Table A.10: Impact of decentralization on types of politicians (estimates of δ)
Alternative definitions dependent variable

	(1)	(2)	(3)	(4)
	Y^a [A]	Y^p [A]	Y^a [B]	Y^p [B]
Panel A: without control variables				
decentr \times FiscAut	0.194*** (0.037)	-0.300*** (0.047)	0.118*** (0.035)	-0.294*** (0.058)
obs.	1,304	1,304	1,304	1,304
Panel B: with control variables				
decentr \times FiscAut	0.195*** (0.040)	-0.301*** (0.042)	0.113*** (0.035)	-0.308*** (0.048)
	1,304	1,304	1,304	1,304
Year FE	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: Y^a [A] = 1 if adm_jobs [A] = 1 and years_pol_exp = 0; Y^a [B] = 1 if adm_jobs [B] = 1 and years_pol_exp = 0; Y^p [A] = 1 if adm_jobs [A] = 0 and years_pol_exp > 0; Y^p [B] = 1 if adm_jobs [B] = 0 and years_pol_exp > 0; adm_jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; adm_jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; years_pol_exp: years of past political experience of the mayor at any institutional level. Independent variables definition: decentr: dummy variable equal to 1 after the introduction of the property tax (ICI); FiscAut: variable that captures the level of fiscal autonomy of the municipality. In this case it is equal to gdp per capita measured in the years 1992-1993. Additional control variables added to the regression but not reported in the Table (Panel B only): municipal population; share of municipal population older than 65 years old; # firms not interacted with decentr, political competition. The variables % college, % referendum and # no profit are time invariant, so they are already captured by municipal fixed effects.

Appendix 2 [For Online Publication]

Derivation of $f(z^j)$.

Consider $x^j = \alpha t \theta^j$ and $y^j = \tau(\alpha) \delta^j$. Clearly, x^j and y^j are also uniformly distributed random variables, with density $\frac{\psi}{\alpha t}$ and $\frac{\psi}{\tau(\alpha)}$, respectively. The extremes of the two variables are: $\underline{y}^j = \tau(\alpha)(-\frac{1}{2\psi} + \bar{\delta}^j)$; $\bar{y}^j = \tau(\alpha)(\frac{1}{2\psi} + \bar{\delta}^j)$, $\underline{x}^j = \alpha t(-\frac{1}{2\psi} + \bar{\theta}^j)$, $\bar{x}^j = \alpha t(\frac{1}{2\psi} + \bar{\theta}^j)$. Let $k = \frac{\psi^2}{\alpha t \tau(\alpha)}$. By the convolution theorem, under A.1, $f(z^j)$ is:

$$\begin{aligned} f(z^j) &= k(z^j - \underline{x}^j - \underline{y}^j), \text{ for } \underline{x}^j + \underline{y}^j \leq z^j \leq \bar{x}^j + \underline{y}^j; \\ f(z^j) &= k(\bar{x}^j - \underline{x}^j) = k \frac{\alpha t}{\psi} = \frac{\psi}{\tau(\alpha)}, \text{ for } \bar{x}^j + \underline{y}^j \leq z^j \leq \underline{x}^j + \bar{y}^j; \\ f(z^j) &= k(\bar{x}^j + \bar{y}^j - z^j), \text{ for } \underline{x}^j + \bar{y}^j \leq z^j \leq \bar{x}^j + \bar{y}^j; \end{aligned} \quad (10)$$

Note that the second part of A1 implies that $E(z^k)$ belongs to the "flat" part of $f(z^j)$, for $j, k = a, p$.

9.1 Derivation of r_1^{j*} .

Rewriting Eq.(4) fully,

$$\begin{aligned} E(R^j) &= \lambda \bar{R} + r_1^j E(z^j) - \lambda \bar{R} (\pi \int_{\underline{z}^j}^{E(z^a) \frac{(1-r_1^{je})}{(1-r_1^j)}} f(z^j) dz^j + \\ &\quad + (1 - \pi) \int_{\underline{z}^j}^{E(z^p) \frac{(1-r_1^{je})}{(1-r_1^j)}} f(z^j) dz^j) \end{aligned}$$

where $\underline{z}^j = \underline{x}^j + \underline{y}^j$. Differentiating for r_1^j ,

$$\begin{aligned} E'(R^j) &= E(z^j) - \lambda \bar{R} (\pi f(E(z^a) \frac{(1-r_1^{je})}{(1-r_1^j)}) E(z^a) \frac{(1-r_1^{je})}{(1-r_1^j)^2} + \\ &\quad + (1 - \pi) f(E(z^p) \frac{(1-r_1^{je})}{(1-r_1^j)}) E(z^p) \frac{(1-r_1^{je})}{(1-r_1^j)^2}) \end{aligned}$$

at the equilibrium, $r_1^j = r_1^{je}$ must hold. Hence:

$$\begin{aligned} E'(R^j) &= E(z^j) - \lambda \bar{R} (\pi f(E(z^a)) E(z^a) \frac{1}{(1-r_1^j)} + \\ &\quad + (1 - \pi) f(E(z^p)) E(z^p) \frac{1}{(1-r_1^j)}) \end{aligned}$$

Recall from Eq. (10) that $f(z^j) = \frac{\psi}{\tau(\alpha)}$, for

$$E(z^j) - \frac{1}{2\psi}(\tau(\alpha) - \alpha t) \leq z \leq E(z^j) + \frac{1}{2\psi}(\tau(\alpha) - \alpha t)$$

Under A.1, this clearly holds for both $z = E(z^a)$ and $z = E(z^p)$.

Substituting for $f(z^j)$ and imposing $E'(R^j) = 0$,

$$(1 - r_1^j)E(z^j) = \frac{\lambda\bar{R}\psi}{\tau(\alpha)}(\pi E(z^a) + (1 - \pi)E(z^p)) = \frac{\lambda\bar{R}\psi}{\tau(\alpha)}E(z^o)$$

that solving for r_1^j , gives

$$r_1^{j*} = 1 - \frac{\lambda\bar{R}\psi}{\tau(\alpha)} \frac{E(z^o)}{E(z^j)}.$$

QED.

Proof of Proposition 1

For the proposed strategies to form an equilibrium, the incumbent j must prefer to play his proposed strategy in the first period rather than deviating immediately and take maximal rents (and not be re-elected). This requires $E(R^{j*}) \geq \bar{R}$ for both $j = a, p$. Let then $E(z^k) = \min(E(z^a); E(z^p))$; the candidate more willing to deviate is then type k . For this candidate not to deviate, it must then hold:

$$E(R^{k*}) = E(z^k) - \frac{\lambda\bar{R}\psi}{\tau}E(z^o) + \lambda\bar{R}\left(\frac{1}{2} - \frac{\psi}{\tau}s [E(z^{-k}) - E(z^k)]\right) \geq \bar{R}$$

where $s = \pi$ if $k = p$ and $s = (1 - \pi)$ otherwise, which can be rewritten as:

$$R' = \frac{E(z^k)}{\left(1 - \frac{\lambda}{2} + \frac{\psi\lambda}{\tau} [E(z^k) + s2(E(z^{-k}) - E(z^k))]\right)} \geq \bar{R} \quad (11)$$

Notice from (5), that (11) also implies $r_1^{k*} > 0$. Hence, as stated in Proposition 1, for $R' > \bar{R}$ we then get an equilibrium. To prove that this equilibrium is also unique, compute $r_1^j(r_1^{je}) = \arg \max E(R^j)$ for an arbitrary value of r_1^{je} and note that $r_1^j(r_1^{je}) = r_1^{je}$ only for $r_1^{j*} = 1 - \frac{\lambda\bar{R}\psi}{\tau} \left[\frac{E(z^o)}{E(z^j)} \right]$. QED.

Proof of Proposition 2

The effect of a small change in t on the expected rents for the two types can be found by differentiating:

- for a -type

$$E(R^a) = E(z^a) + \lambda\bar{R}\frac{\psi}{\tau}((1 - 2\pi)E(z^a) - 2(1 - \pi)E(z^p)) + \lambda\bar{R}\frac{1}{2}$$

Solving:

$$\frac{d}{dt}E(R^a) = \alpha(\bar{\theta}^a - \bar{\delta}^a) + \lambda\bar{R}\frac{\psi^a}{\tau^2}((1-2\pi)(\bar{\theta}^a - \bar{\theta}^p) - \bar{\theta}^p)$$

which is positive if

$$\bar{\theta}^a(1 + \lambda\bar{R}\frac{\psi}{\tau^2}(1-2\pi)) > \bar{\delta}^a + \lambda\bar{R}\frac{\psi}{\tau^2}2(1-\pi)\bar{\theta}^p.$$

This holds if

$$\bar{\theta}^a \geq \frac{2(1-\pi)}{(1-2\pi)}\bar{\theta}^p.$$

The worst possible case is $\pi = 1$. But even in this case $\frac{d}{dt}E(R^a) > 0$ if

$$\bar{\theta}^a \geq \frac{\bar{\delta}^a}{1 - \lambda\bar{R}\frac{\psi}{\tau^2}} = \frac{\bar{\delta}^a \tau}{(1-at) - \lambda\bar{R}\frac{\psi}{\tau}} = \theta^*.$$

Consider now

$$\frac{d^2}{dt da}E(R^a) = (\bar{\theta}^a - \bar{\delta}^a) + \lambda\bar{R}\frac{\psi}{\tau^2}\frac{(1+at)}{(1-at)}((1-2\pi)(\bar{\theta}^a - \bar{\theta}^p) - \bar{\theta}^p)$$

which is positive if

$$\bar{\theta}^a(1 + \lambda\bar{R}\frac{\psi}{\tau^2}\frac{(1+at)}{(1-at)}((1-2\pi))) > \bar{\delta}^a + \lambda\bar{R}\frac{\psi}{\tau^2}2(1-\pi)\bar{\theta}^p.$$

This holds if

$$\bar{\theta}^a \geq \frac{2(1-\pi)}{(1-2\pi)}\bar{\theta}^p.$$

In the worst possible case $\pi = 1$, $\frac{d^2}{dt da}E(R^a) > 0$ if

$$\theta^a \geq \frac{\delta^a}{1 - \lambda\bar{R}\frac{\psi}{\tau^2}\frac{(1+at)}{(1-at)}} = \theta^{**}.$$

where $\theta^{**} > \theta^*$.

- For p -type:

$$E(R^p) = E(z^p) + \lambda\bar{R}\frac{\psi}{\tau}(-2\pi E(z^a) - (1-2\pi)E(z^p)) + \lambda\bar{R}\frac{1}{2}$$

$$\frac{d}{dt}E(R^p) = \alpha(\bar{\theta}^p - \bar{\delta}^p) + \lambda\bar{R}\frac{\psi^a}{\tau^2}(2\pi(\bar{\theta}^p - \bar{\theta}^a) - \bar{\theta}^p) < 0.$$

and

$$\frac{d^2}{dt da}E(R^p) = (\bar{\theta}^p - \bar{\delta}^p) + \lambda\bar{R}\frac{\psi}{\tau^2}(2\pi(\bar{\theta}^p - \bar{\theta}^a) - \bar{\theta}^p) + 2\lambda\bar{R}\frac{\psi at}{\tau^3}(2\pi(\bar{\theta}^p - \bar{\theta}^a) - \bar{\theta}^p) < 0$$

Finally:

$$\frac{d(E(R^a) - E(R^p))}{dt} = \alpha(\bar{\theta}^a - \bar{\delta}^a + \bar{\delta}^p - \bar{\theta}^p) + \lambda\bar{R}\frac{\psi a}{\tau^2}(\bar{\theta}^a - \bar{\theta}^p) > 0;$$

$$\frac{d^2(E(R^a) - E(R^p))}{dt d\alpha} = \bar{\theta}^a - \bar{\delta}^a - \bar{\theta}^p + \bar{\delta}^p + \lambda\bar{R}\frac{\psi}{\tau^2}\frac{(1 + \alpha t)}{(1 - \alpha t)}(\bar{\theta}^a - \bar{\theta}^p) > 0.$$

QED.

Proof of Proposition 3

Writing it in full, the expected utility of the consumer under the two types of incumbent can be written:

$$\begin{aligned} U^a = & \frac{\lambda\bar{R}\psi}{\tau}E(z^o) - \lambda\bar{R} + \lambda(\pi(\frac{1}{2}(E(z^a) + E(z^a|z^a \geq E(z^a))) + \\ & + \lambda(1 - \pi)((\frac{1}{2} + \frac{\psi}{\tau}h)E(z^a|z^a \geq E(z^p)) + (\frac{1}{2} - \frac{\psi}{\tau}h)E(z^p)) \end{aligned} \quad (12)$$

$$\begin{aligned} U^p = & \frac{\lambda\bar{R}\psi}{\tau}E(z^o) - \lambda\bar{R} + \lambda((1 - \pi)(\frac{1}{2}(E(z^p) + \\ & + E(z^p|z^p \geq E(z^p))) + \lambda\pi((\frac{1}{2} - \frac{\psi}{\tau}h)E(z^p|z^p \geq E(z^a)) + \\ & + (\frac{1}{2} + \frac{\psi}{\tau}h)E(z^a)) \end{aligned} \quad (13)$$

where $h = E(z^a) - E(z^p)$. Invoking (10), and solving the integrals we obtain:

$$E(z^a|z^a \geq E(z^a)) = E(z^a) + T$$

$$E(z^a|z^a \geq E(z^p)) = \frac{\frac{1}{2}(E(z^a) + T + \frac{\psi}{\tau}hy)}{(\frac{1}{2} + \frac{\psi}{\tau}h)}$$

$$E(z^p|z^p \geq E(z^p)) = E(z^p) + T$$

$$E(z^p|z^p \geq E(z^a)) = \frac{\frac{1}{2}(E(z^p) + T - \frac{\psi}{\tau}hy)}{(\frac{1}{2} - \frac{\psi}{\tau}h)}$$

where $T = \frac{4\tau^2 + 1 - 2\tau}{12\tau\psi} > 0$ and $y = E(z^a) + E(z^p)$.

Substituting in (12) and (13) and simplifying:

$$U^a = \frac{\lambda\bar{R}\psi}{\tau}E(z^o) - \lambda\bar{R} + \lambda\frac{T}{2} + \lambda\pi E(z^a) + \lambda(1 - \pi)(\frac{1}{2}(E(z^a) + \frac{\psi}{\tau}hy) + (\frac{1}{2} - \frac{\psi}{\tau}h)E(z^p))$$

$$U^p = \frac{\lambda \bar{R} \psi}{\tau} E(z^o) - \lambda \bar{R} + \lambda \frac{T}{2} + \lambda(1 - \pi)E(z^p) + \lambda \pi \left(\frac{1}{2} (E(z^p) - \frac{\psi}{\tau} h y) + \left(\frac{1}{2} + \frac{\psi}{\tau} h \right) E(z^a) \right)$$

Now let $U = \pi U^a + (1 - \pi) U^p$. Differentiating U with respect to t and recalling that $\tau = 1 - \alpha t$:

$$\begin{aligned} \partial U / \partial t &= \frac{\lambda \bar{R} \psi}{\tau^2} \alpha (\pi \bar{\theta}^a + (1 - \pi) \bar{\theta}^p) + \lambda \alpha \left[\frac{1 - 4\tau^2}{24\tau^2 \psi} \right] + \\ &\quad + \lambda (\pi^2 dE(z^a)/dt + (1 - \pi)^2 dE(z^p)/dt) + \\ &\quad + \lambda \pi (1 - \pi) (dy/dt + \frac{\psi}{\tau^2} \alpha h (2(\bar{\theta}^a - \bar{\theta}^p) - h)) \end{aligned}$$

The first term is surely positive, the second is surely negative (as $\tau > \frac{1}{2}$), the third depends on π , and the fourth can have either sign. Note that for both $\pi = 0$ and $\pi = 1$, the fourth term is zero. Considering these two extreme cases first, it is easy to establish that there exist $\bar{\theta}^{a*} > \bar{\theta}^{p*} > 0$ such that $\pi \rightarrow 0$ and $\bar{\theta}^p \leq \bar{\theta}^{p*}$ implies $\partial U / \partial t < 0$ and $\pi \rightarrow 1$ and $\bar{\theta}^a \geq \bar{\theta}^{a*}$ implies $\partial U / \partial t > 0$. This proves (i). Differentiating $\partial U / \partial t$ with respect to π :

$$\begin{aligned} \partial^2 U / \partial t \partial \pi &= \frac{\lambda \bar{R} \psi}{\tau^2} \alpha (\bar{\theta}^a - \bar{\theta}^p) + 2\lambda (\pi E'(z^a) - (1 - \pi) E'(z^p)) + \\ &\quad + \lambda (1 - 2\pi) (dy/dt + \frac{\psi}{\tau^2} \alpha h (2(\bar{\theta}^a - \bar{\theta}^p) - h)) \end{aligned}$$

The first two terms are strictly positive; the third is generally uncertain. But going through element by element, it is easy to check that the third term is dominated by the first two terms for any value of π . Hence, $\partial^2 U / \partial t \partial \pi > 0$. Together with (i) this implies that for any value of α there exists a unique value of π , $\underline{\pi}(\alpha)$ such that $\partial U / \partial t = 0$ and that $\pi \geq \underline{\pi}(\alpha)$ implies $\partial U / \partial t \geq 0$. This proves (ii). Finally, evaluating $\partial U / \partial t$ at $\underline{\pi}(\alpha)$, and totally differentiating $\frac{d\underline{\pi}}{d\alpha} = -\frac{\partial^2 U / \partial t \partial \alpha}{\partial^2 U / \partial t \partial \pi}$. The denominator is strictly positive, so that $sign(\frac{d\underline{\pi}}{d\alpha}) = -sign(\partial^2 U / \partial t \partial \alpha)$. Differentiating $\partial U / \partial t$ with respect to α and exploiting the fact that at $\underline{\pi}(\alpha)$, $\partial U / \partial t = 0$, it can be shown that:

$$\begin{aligned} \partial^2 U / \partial t \partial \alpha &= \frac{\lambda \bar{R} \psi \alpha}{\tau^3} 2t (\underline{\pi} \bar{\theta}^a + (1 - \underline{\pi}) \bar{\theta}^p) + \lambda \left(\frac{\alpha t}{12\tau^3 \psi} \right) + \\ &\quad + \lambda \underline{\pi} (1 - \underline{\pi}) 2\alpha \frac{\psi t}{\tau^3} [(\bar{\theta}^a - \bar{\theta}^p)^2] > 0 \end{aligned}$$

It follows that $\frac{d\underline{\pi}}{d\alpha} < 0$. This proves (iii). QED.

Derivation of π^*

Consider again the expected utility of the two types of politicians

$$E(R^a) = E(z^a) - \frac{\lambda\bar{R}\psi}{\tau}E(z^o) + \lambda\bar{R}\left(\frac{1}{2} + \frac{\psi}{\tau}(1 - \pi)[E(z^a) - E(z^p)]\right)$$

$$E(R^p) = E(z^p) - \frac{\lambda\bar{R}\psi}{\tau}E(z^o) + \lambda\bar{R}\left(\frac{1}{2} - \pi\frac{\psi}{\tau}[E(z^a) - E(z^p)]\right)$$

and recall from the main text that endogenous candidacy implies $\pi = \frac{E(R^a)}{E(R^p) + E(R^a)}$.
Rewriting and simplifying:

$$A = a + ak(1 - 2\pi) - 2pk(1 - \pi) + \frac{1}{2}\lambda\bar{R}$$

$$B = p + ak(-2\pi) - pk(1 - 2\pi) + \frac{1}{2}\lambda\bar{R}$$

where $A = E(R^a)$, $B = E(R^p)$, $E(z^a) = a$, $E(z^p) = p$; $k = \frac{\lambda\bar{R}\psi}{\tau}$. It follows that $A - B = (a - p)(1 + k)$. Note that:

$$2\pi - 1 = \frac{(a - p)(1 + k)}{A + B} \Rightarrow \pi = \frac{1}{2} + \frac{(a - p)(1 + k)}{2(A + B)}$$

Summing and substituting we obtain:

$$A + B = a + p + (a - p)k\left(-1 - \frac{2(a - p)(1 + k)}{(A + B)}\right) - 2pk + \lambda\bar{R};$$

which can be rewritten as:

$$(A + B)^2 = (\lambda\bar{R} + (a + p)(1 - k))(A + B) - 2(a - p)^2k(1 + k);$$

Let $A + B = x$; $\lambda\bar{R} + (a + p)(1 - k) = b$; $2(a - p)^2k(1 + k) = c$, leading to $x^2 - bx + c = 0$. Solving, the two roots are $\frac{1}{2}b + \frac{1}{2}\sqrt{b^2 - 4c}$; $\frac{1}{2}b - \frac{1}{2}\sqrt{b^2 - 4c}$. The equation admits real solutions if $b^2 \geq 4c$. This is certainly the case for $|a - p| \leq \frac{b}{2\sqrt{2(k+k^2)}} = Q$. Note that Q is decreasing in k and that for $k \rightarrow 1$, $Q \rightarrow \frac{\tau}{4\psi}$. As $k < 1$, $|a - p| \leq Q$ is then a very mild condition, that is already implied by A.1. Note further that to make economic sense $|\frac{(a-p)(1+k)}{2(A+B)}| \leq \frac{1}{2}$; this would certainly be violated by the negative root for $a \rightarrow p$. Thus, the only economic significant solution is represented by the positive root. Under the positive root, $A + P = \frac{n(b + \sqrt{b^2 - 4c})}{2\bar{w}(1 + \lambda)}$. Substituting, this gives equation (6) into the main text. QED.

Proof of Proposition 4

Differentiating the numerator and the denominator of (6) for t and imposing $\tau = 1 - \alpha t$, we get:

$$\frac{d}{dt}(E(z^a) - E(z^p))\left(1 + \frac{\lambda\bar{R}\psi}{\tau}\right) = \alpha((\bar{\theta}^a - \bar{\theta}^p)\left(1 + \frac{\lambda\bar{R}\psi}{\tau^2}\right) + (\bar{\delta}^p - \bar{\delta}^a)) > 0$$

$$\frac{d}{dt}b = \alpha((\bar{\theta}^a + \bar{\theta}^p)(1 - \frac{\lambda\bar{R}\psi}{\tau^2}) - (\bar{\delta}^a + \bar{\delta}^p))$$

$$\frac{d}{dt}c = 2\alpha(E(z^a) - E(z^p))(\frac{\lambda\bar{R}\psi}{\tau}) \left[(\bar{\theta}^a - \bar{\theta}^p)(1 + \frac{1}{\tau} + 2(\frac{\lambda\bar{R}\psi}{\tau^2})) + (\bar{\delta}^p - \bar{\delta}^a) \right] \geq (<)0$$

for $E(z^a) \geq (<)E(z^p)$. Notice that $(1 - \frac{\lambda\bar{R}\psi}{\tau^2}) = \frac{1}{\tau}(1 - \alpha t - \frac{\lambda\bar{R}\psi}{\tau})$ is of uncertain sign and could be negative, implying $\frac{d}{dt}b < 0$. But even if positive it is a small number and $\frac{d}{dt}b$ would still be negative provided that $\bar{\theta}^a - \bar{\delta}^a$ is not much larger than $\bar{\delta}^p - \bar{\theta}^p$. Assuming this not to be the case, $\frac{d}{dt}b \leq 0$. Under this mild condition, it also follows $\frac{d}{dt}(b + (b^2 - 4c)^{\frac{1}{2}}) < 0$ for $E(z^a) \geq E(z^p)$. The sign of $\frac{d}{dt}(b + (b^2 - 4c)^{\frac{1}{2}})$ is uncertain for $E(z^a) < E(z^p)$ as c' becomes negative. Now let simplify the notation by writing:

$$\pi^* - \frac{1}{2} = \frac{(E(z^a) - E(z^p))(1 + \frac{\lambda\bar{R}\psi}{\tau})}{b + (b^2 - 4c)^{\frac{1}{2}}} = \frac{m}{s}$$

It follows:

$$\begin{aligned} \frac{d}{dt}(\pi^* - \frac{1}{2}) &= \frac{d}{dt} \frac{m}{s} = m's^{-1} - ms^{-2}s' = \frac{1}{s^2}(m's - ms') = \\ &= \frac{1}{s}(m' - \frac{m}{s}s') = \frac{1}{s}(m' - (\pi^* - \frac{1}{2})s') \end{aligned}$$

For $\frac{1}{2} \leq \pi^*$ implies $\frac{d}{dt}\pi^* > 0$, as $m' > 0$ and $s' < 0$. For $\frac{1}{2} > \pi^*$, the sign of s' becomes uncertain (as $c' < 0$). If s' is still negative (the size of the politician falls following the reform even in municipalities with $\alpha < \alpha^*$), which is certainly the case for π^* close to $\frac{1}{2}$, the sign of $\frac{d}{dt}\pi^*$ becomes uncertain. By continuity, the sign of m' however dominates the sign of $\frac{m}{s}s'$ for $\frac{m}{s}$ close to zero, implying $\frac{d}{dt}\pi^* > 0$. But for other values of $\frac{1}{2} > \pi^*$, $\frac{d}{dt}\pi^*$ might become negative if the sign of s' is still negative. QED