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

Massimo Bordignon†
Università Cattolica del Sacro Cuore, Milan & CESifo

Matteo Gamalerio
Universitat de Barcelona, Institut d'Economia de Barcelona (IEB)

Gilberto Turati
Università Cattolica del Sacro Cuore, Rome

February 7, 2020

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 the editor, two anonymous referees, Fernanda Brollo, 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 Sobbrio, 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, Università Cattolica del Sacro Cuore, 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 a 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 Mokerjee, 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, 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; Turati, Montolio, and Piacenza, 2017). In addition, 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, Rizzo, and 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 fewer incentives in controlling how much money is spent if a large part of it does not come from their own pockets.[2] 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.

[1] Treisman (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 to support decentralization.

[2] These 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.

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

In this paper we discuss an additional argument, which links the level of fiscal autonomy to 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 depends on the characteristics of a jurisdiction. In particular, in a decentralized setting where most of the resources come from the center, the main task of a local politician is to make sure that these resources keep flowing to the local community (Borck and Owings, 2003; Dalle Nogare and Kauder, 2017). 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 the 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 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 center 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., Galasso and Nannicini, 2011; Gagliarducci and Nannicini, 2013; Gamalerio, 2019), we use the 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 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[4] (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)[5] 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" (Besley and Smart, 2007).

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* for administrative/political skills and not for 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 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 the local level supports

[4]As explained below, past working experience seems to capture in a better way the level of administrative skills of the mayors, compared to other personal characteristics such as education, gender, and age.

[5]In most of the empirical analysis we use the pre-reform GDP per-capita as a proxy for fiscal autonomy. This choice is because, the more direct measures of fiscal autonomy that are available are either measured in the post-treatment period, and thus they are potentially endogenous, or they come from a simulation and not real data. As described in section 5.3, our results are unchanged if we use these alternative measures for fiscal autonomy.

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 occurred in Italy at the beginning of the '90s following the "Clean Hands" trial, which leads 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 controlling 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), for the costs of electoral campaigns, and for potential differences between the Centre-North and the South of Italy.

Our findings have important implications for the debate on fiscal federalism, which we discuss in the conclusions. Clearly, not all recipes are adaptable to all circumstances: fiscal decentralization could 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 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 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 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 occur even with an unchanged budget and with politicians of different skills having the same preferences for rents. Additionally, 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 consider also the case of an increase in local taxation accompanied by an offsetting reduction 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.

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 two 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 of 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 period.[6] Local politicians only care about collecting as many rents as possible from the 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.

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 the detriment of voters. More specifically, in period 1, when an incumbent politician of type j is in charge, we assume:

$$g_1^j = \left[\alpha t \theta^j + \tau(\alpha)\delta^j\right](1 - r_1^j) \tag{2}$$

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 $\overline{\alpha} > \alpha > \underline{\alpha} > 0$. $\tau(\alpha) > 0$ 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. $\tau'(\alpha) < 0$).[7]

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

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

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 (Borck and Owings, 2003; Dalle Nogare and Kauder, 2017). Similarly, a politician with a 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, through which it is then possible to generate a higher level of public good provision.[8] Notice from Eq. (2) that we also assume that a politician can divert to his advantage the extra resources that he 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 $\overline{\theta}^j > 0$, $\overline{\delta}^j > 0$, respectively. In keeping with the discussion above, we assume $\overline{\theta}^a > \overline{\delta}^a, \overline{\delta}^p > \overline{\theta}^p$, $\overline{\theta}^a > \overline{\theta}^p$, $\overline{\delta}^a < \overline{\delta}^p$. 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 C). Specifically, in the following we assume:

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

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.

[8] From this point of view, θ^j can be interpreted as the ability to raise revenues from local taxes. In practice, a politician with a high level of administrative skills is a good collector of taxes. The revenues from local taxes can then be used to provide higher levels of public goods. A potential criticism of the intuition behind Eq. (2) is that administrative skills do not affect the whole municipal budget, including transfers from the central government. However, a more general formulation like $g_1^j = \theta^j[\alpha t + \tau(\alpha)\delta^j](1 - r_1^j)$ would be analytically intractable. For this reason, we have decided to keep the simpler version reported in Eq. (2).

[9] Details 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 signaling 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 a few months before the elections. However, as will become clear as we proceed, relaxing this assumption would not affect much our results.[10]

To solve the model, we work backward. In period 2, as there is no future ahead, whoever is in charge takes maximal rents, $R_2^k = \overline{r} z^k$, where $\overline{r} < 1$ 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., $R_2^k = \overline{R} > 0$ for $k = a, p$ and $s = 1, 2$.[11] In the second period, the utility of the voter is then $z^k - \overline{R}$. 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:

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

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

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

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:

$$E(R^j) = r_1^j E(z^j) + \lambda \overline{R}\pi \left[1 - prob\left(z^j \leq E(z^a)\frac{(1-r_1^{je})}{(1-r_1^j)}\right)\right] + \tag{4}$$
$$+\lambda \overline{R}(1-\pi)\left[1 - prob\left(z^j \leq E(z^p)\frac{(1-r_1^{je})}{(1-r_1^j)}\right)\right]$$

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 rate (see Appendix C for a formal proof.):

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

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 $\overline{R}$ in the first period (and not be re-elected in the second). As shown in Appendix C, this translates in an upper limit for $\overline{R}$, $\overline{R} < R'$; 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 $r_1^{j*}E(z^j) = E(z^j) - \frac{\lambda \overline{R}\psi}{\tau}E(z^o)$; hence, the (expected) level of the public good in the first period is just $g_1^* = \frac{\lambda \overline{R}\psi}{\tau}E(z^o)$ 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 equilibrium.[12] 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 $\frac{1}{2} + \frac{\psi}{\tau}\left[E(z^j) - E(z^k)\right]$, $j,k = a,p$, $j \neq k$ 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.

[12]As shown, expected rents are instead decreasing in $\lambda\overline{R}$ (a larger $\lambda\overline{R}$ 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}).

The strategies of the voter at the proposed equilibrium are straightforward. The voter sets 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 $R' > \overline{R}$. Then there exists a unique equilibrium where the voter sets up a threshold for g_1, $g_1^{jk} = \frac{\lambda \overline{R} \psi}{\tau} E(z^o) \frac{E(z^k)}{E(z^j)}$ 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 $r_1^{j*} = 1 - \frac{\lambda \overline{R} \psi}{\tau} \left[\frac{E(z^o)}{E(z^j)} \right]$, 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 C. ■

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 center.[13] 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 $\alpha^* = \frac{\overline{\delta}^p - \overline{\delta}^a}{t(\overline{\theta}^a - \overline{\theta}^p + \overline{\delta}^p - \overline{\delta}^a)}$, provided that $\overline{\alpha} > \alpha^* > \underline{\alpha}$. 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.

[13] Notice 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 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 politicians:[14]

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* $\overline{\theta}^a \geq \theta^* > 0$; *then* $\partial E(R^a)/\partial t > 0$. *Suppose* $\overline{\theta}^a \geq \theta^{**} > \theta^*$, *then* $\partial^2 E(R^a)/\partial\alpha\partial t > 0$.

Proof. See Appendix C. ■

To provide an intuition for Proposition 2, note that a change in t, matched by 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 $\overline{\theta}^j$ and $\overline{\delta}^j$, 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 $\overline{\theta}^a$ 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 relative 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* $\overline{\theta}^a > \overline{\theta}^{a^*} > \overline{\theta}^{p*} > \overline{\theta}^p > 0$. *Then, (i)* $\partial U(\alpha)/\partial t < 0$ *for* $\pi \to 0$; $\partial U(\alpha)/\partial t > 0$ *for* $\pi \to 1$.*(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 C. ■

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

[14] In 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 a 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 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 $\{0, \overline{w}\}$. 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 $\frac{1}{\overline{w}}\frac{E(R^j)}{(1+\lambda)}$, 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 $J = \frac{1}{\overline{w}}\frac{E(R^j)}{(1+\lambda)}n$, 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 C shows, solving the resulting system of simultaneous equations and assuming an interior solution, this equilibrium share can be computed as:

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

where $c = \frac{2\overline{w}(1+\lambda)}{n} > 0$. 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-type.[15] 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 C. ■

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 C, 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 dominate 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 reinforces 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

[15] Private 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 $\overline{w}^a > \overline{w}^p$. 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 are available on request from the authors.

it increases the utility of voters for a given π; and indirectly because it also increases π. On the 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 in 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 supporting 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 old proportional electoral rule in favor of systems that made elected politicians more directly accountable to citizens.[16] 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 the local level, reforming a system that was traditionally based on transfers and tied grants.[17] This began what is known as the "decentralization" period in Italy, that culminated in 2001 Constitutional reform (see Ambrosanio et al., 2015).

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

[16]Specifically, 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).

[17]Specifically, 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).

for parties' representatives to elect the city council, which 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 council.[18] 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 only. 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 rules. 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 allowances. Tax 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. 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 described by Bordignon et al. (2017), the PIT surcharge represented a way to provide to the municipal governments a more elastic source of fiscal revenues, different from the property tax (ICI), whose tax base was more rigid. The PIT surcharge further increased the level of tax autonomy of the municipalities, even though to a lower extent compared to the property tax (ICI). Differently from the property tax (ICI), the PIT surcharge was not followed by an offsetting cut in fiscal transfers and its introduction was not mandatory for the municipal governments, which could freely decide whether to apply it or not within the limit of the 0.5% of the taxable income. In 2003, the central government froze this autonomy in setting the surcharge rate, forcing the municipalities to keep the rate set in the previous year.

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. Following these two tax decentralization reforms, municipalities started to enjoy a great level of fiscal autonomy, with the property tax (ICI) and the PIT surcharge representing the core of this autonomy. In 2000, for example, the property tax (ICI) represented about 50% of total municipal tax revenues and the PIT surcharge approximately 8% of total municipal tax revenues (Bordignon et al.

[18]There 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.

2003; Bordignon et al., 2017). 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 grants covered on average about 60-70% of total municipal expenditure (ISTAT, *I Bilanci consuntivi delle amministrazioni comunali*, June 2012).

4 The empirical strategy

The Italian case offers an interesting testing grounds for our theory. The main challenge to attribute a causal impact to the tax reform is that the reform in the electoral rule and the introduction of the property tax happened in the same year, 1993, thus potentially confounding the two effects; we discuss below several different ways in which we address this problem. 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 $decentralization_t$ identifies the post-treatment years (i.e., all the years following the introduction of the ICI property tax[19]), while $fiscal\ autonomy_i$ is our measure of fiscal autonomy (see the definition below). The interaction term $decentralization_t \times fiscal\ autonomy_i$ represents our treatment variable, allowing the impact of the tax decentralization reform to differ across rich and poor cities.[20] 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:

$$Y_{it}^{j} = A_i + B_t + \gamma X_{it} + \delta decentralization_t \times fiscal\ autonomy_i + \varepsilon_{it} \tag{7}$$

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.

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: $decentralization1_t$ identifies the years between the introduction of ICI and the introduction of the PIT surcharge (1993-1998), while $decentralization2_t$ takes value one after the introduction of the PIT surcharge (1999-2002).[21] The important feature of this second approach is that, while $decentralization1_t$ identifies the shock stemming from the new electoral rules and the new property tax, $decentralization2_t$ iden-

[19] We 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.

[20] We follow a multivalued treatment approach; see, e.g., Imbens and Wooldridge, 2009.

[21] Our 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.

tifies the *additional* shock on fiscal autonomy at *unchanged electoral rules*, so that any further differential effect stemming from the treatment variable $fiscal\ autonomy_i$ could be attributed solely to tax decentralization in this second period. This offers the first attempt to separate the impact of tax decentralization from the impact of the new electoral rules. The second model can be written as:

$$Y_{it}^{j} = A_i + B_t + \gamma X_{it} + \tag{8}$$
$$+\delta_1 decentralization1_t \times fiscal\ autonomy_i + \delta_2 decentralization2_t \times fiscal\ autonomy_i + \varepsilon_{it}$$

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 *decentralization*:

$$Y_{it}^{j} = A_i + B_t + \gamma X_{it} + \tag{9}$$
$$+\delta decentralization_t \times fiscal\ autonomy_i + \theta decentralization_t \times X_i + \varepsilon_{it}$$

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

5.1.1 The sample

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 Affairs (*Anagrafe degli Amministratori Locali*). 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 Table A.1.[22] 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,

[22] In Table A.2, we show how the municipalities in our sample differ along various dimensions, distinguishing between municipalities above and below the median GDP. The evidence in Table A.2 shows that the two groups of municipalities differ along almost all the dimensions considered. For this reason, as robustness checks in the analysis below, we show that our results are robust to the inclusion of these control variables.

such as internet web pages and newspapers' archives, 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).

5.1.2 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 *administrative 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 (*administrative 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, veterinarians, dentists, traders, biologists, university professors, surveyors, business consultants, and pharmacists. The second dummy variable (*administrative jobs*$[B]$) is defined, somewhat more arbitrarily, by excluding from *administrative jobs*$[A]$ a number of professions whose level of administrative skills is more disputable.[23] Thus, in this final and more restricted definition, we only remain with school managers, entrepreneurs, directors, engineers and architects, traders, surveyors, and business consultants.

In building our measure for "administrative" skills, we focus on past working experience because we think that this variable captures in a better way the skills suggested by the theoretical model. Evidence from our data and the characteristics of the Italian labor market provides

[23]For 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 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.

also good justifications for the use of past working experience rather than other personal characteristics. First, while the literature has shown that other personal characteristics such as education (Besley et al., 2011; Martinez-Bravo, 2017), gender (Brollo and Troiano, 2016) and age (Alesina, Cassidy, and Troiano, 2018) may have an impact on policy outcomes, as shown in Table A.3, a simple correlation analysis from our data suggests that other personal characteristics of the mayors only weakly correlate with the ISFOL index measuring the importance of the administrative skills associated with the past occupation of the mayor. Second, according to Eurostat data, Italy was a country where only approximately 25 % of the managers in the labor market had a university degree, one of the lowest shares across European countries and certainly smaller than the one of other comparable countries like, for instance, France, where the share was above 40 %. This evidence suggests that, in the Italian context and the period studied, the correlation between education and administrative skills was low. Hence, while education may be a more common proxy for ability than past working experience, in this case, it seems less appropriate to capture the "administrative" skills of the Italian mayors.

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.[24] We then define the variable % *political experience* 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, % *political experience* 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 are 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 % *political experience*, i.e., the number of years of previous political experience (*years political experience*) regardless of the age of mayor. The intuition behind the use of past political experience as a proxy for "political" skills is that mayors who have spent a longer period in politics are powerful people and may have many connections to other politicians and other levels of government, independently of their political orientation or the political ideology of the higher levels of government. These connections should enable them to generate benefits for their municipality, including the possibility to attract grants from the central government or the regional government (Carozzi and Repetto, 2016).[25] [26]

[24]We 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.

[25]In Table A.4, using data for the period for which municipalities balance sheets are available (i.e. 2000-2010), we show that this is the case, given that both % *political experience* and *years political experience* positively correlate with the grants received by the municipalities from the central and regional governments. This positive correlation represents a supportive argument for the choice of our proxy for "political" skills.

[26]The reason for using political experience as a proxy for "political" skills rather than other measures such alignment is that we expect the potential beneficial effect of political experience on grants to be more stable and long-lasting over time. While the literature has shown that the alignment status generates political favoritism by part of the central government (Bracco et al., 2015; Brollo and Nannicini, 2012; Curto, Solé Ollé, and Sorribas, 2018), this favoritism may not be stable over time, especially in a country like Italy where political crisis are frequent and where the color of the national government may change quickly even after new elections. For this reason, we may expect a weaker and less precise effect of tax decentralization on the alignment of mayors with higher levels of government.

Descriptive statistics for our outcome variables are in Table A.1. 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). The average years of political experience are about 8, with a range from 0 to 56.

5.1.3 The treatment variable

We interact dummy variables identifying post-treatment years with the variable $fiscal\ autonomy$, a measure of fiscal autonomy. We proxy the fiscal capacity of a municipality with per capita GDP in the province[27] in which the city is located.[28] This is highly correlated with the tax base of both the property tax and the PIT, and it is less subject to endogeneity concerns. 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 euro[29]; this variability will be helpful in identifying our coefficients of interest.

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). Panel A of 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 columns (1)-(3). All the estimated coefficients take up the expected positive signs and are statistically significant at the usual confidence levels. 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 ($administrative\ index$) of approximately 5%, and an increase in the proportion of mayors coming from jobs characterized by high administrative skills between 9% points ($administrative\ jobs[B]$) and 6% points ($administrative\ jobs[A]$). In columns (4)-(5), we report the results for the proportion of p-type politicians. In line with our theoretical predictions, 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 a city with a one standard deviation higher GDP per capita, after the electoral

[27]Provinces 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 a region) and each of the city in our sample it is the administrative capital of one of this province.

[28]Data are provided by the "IstitutoTagliacarne" (http://www.tagliacarne.it/). We use GDP measured at the provincial level because data at the municipal level are not available for all years.

[29]In 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.

reform, the percentage of past working years that mayors had dedicated to politics before being elected (% *political experience*) fell by approximately 3.4% points, while average past political experience (*years political experience*) declined by about 1 year.[30]

In Panel B of Table 1, we report the results obtained by running the second model (8), which separates the effect of the two tax reforms and allows us to test the common trend assumption (see below). The raw effects in Panel A of Table 1 are largely confirmed, in terms of the 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).[31] 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 increases 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 three different strategies. First, we add anticipatory effects to model (8) interacting our measure for fiscal autonomy ($fiscal\ autonomy_i$) with the dummy variable *all 4 years before*, which takes value 1 for the four years before the introduction of the property tax. We select four years before the reform because for most municipalities we observe at most five years in the pre-treatment period, and thus we use the first year of observation as omitted category. In addition, selecting four years enables us to balance *all 4 years before* with the post-reform dummy $decentralization1_t$ and $decentralization2_t$. Results for this exercise are reported in columns (1), (3), and (5) of Table 2 for administrative skills, and in columns (7) and (9) of Table 2 for political skills. As shown in Table 2 and Figure 2, the coefficients for the anticipatory

[30] In Panel B of Table B.5, we repeat the analysis using alternative measures for "political" skills, such as alignment with the central government and/or the regional government (i.e. the two main levels of government from which municipalities receive grants), and a dummy variable equal to 1 for independent mayors supported by local parties, i.e. "Civic Lists" (see Gamalerio, 2020, for more information). Even though the coefficients for the alignment status are not precisely estimated, the results go in the expected direction, with a reduction in the share of aligned mayors and an increase in the share of independent mayors in richer municipalities compared to the poorer ones.

[31] We could expect a bigger point estimate associated with the 1999 reform, as it potentially captures the effects of both the 1993 and the 1999 reforms. Panel B of Table 1 shows that this is not the case. There are a few reasons why we could expect a smaller impact of the PIT surcharge compared to the property tax (ICI). First, the 1999 reform represented a smaller reform, through which the municipalities could freely decide whether to apply the PIT surcharge or not, and the PIT surcharge was not followed by an offsetting reduction in grants. Second, for most of the municipalities, the introduction of the PIT surcharge in 1999 coincided with the second municipal election after the 1993 reform. Considering that many new mayors were elected immediately after 1993, around 1999 most of these new mayors were re-running again as mayoral candidates. Hence, many municipalities did not change the mayor in the period around the introduction of the PIT surcharge. For example, immediately after 1993, only 29 % of the mayors had previous experience as a mayor. This percentage rose to 50 % in the years around 1999. To take into account this fact, in Table B.6, we repeat the analysis distinguishing between municipalities that changed the mayor following the introduction of the PIT surcharge and municipalities that did not. Even though the difference between the two coefficients is not always statistically different from zero, consistent with this intuition, we find a bigger coefficient associated with the 1999 reform for municipalities that changed the mayor in the years around 1999, compared to the coefficient associated to the 1993 reform. The same result does not apply for municipalities that did not change the mayor around 1999.

effects are always insignificant, suggesting that there is no evidence of differential trends in pre-treatment years. Second, we repeat the same exercise splitting the dummy variable *all 4 years before* in four different dummy variables, each for every year before the introduction of the property tax. We interact the four dummy variables with our measure for fiscal autonomy ($fiscal\ autonomy_i$). Columns (2), (4) and (6) of Table 2 and columns (8) and (10) of Table 2 report the results of this exercise. As we can see, none of the coefficients of the four interaction terms is statistically different from zero, suggesting that the municipalities in our sample were not following differential trends in the pre-treatment years.[32]

5.2.2 Discussion: a causal effect? Alternative explanations

Our evidence shows that after the reforms, we observe a differential change between rich and poor cities in the ruling class at the municipal level. 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 some 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 of this problem. Hence, what we do in the following is to carefully review some 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 One of the main threats 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 $decentralization2$ and $fiscal\ autonomy$ in Tables 1 and 2, 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 politicians with different political

[32]Table B.1 reports the results of the same pre-trends robustness check run keeping only the sub-sample of municipalities that change at least one mayor in the period before the introduction of the property tax. As shown, even in this more stringent robustness check we do not find any evidence of differential trends in the pre-treatment period.

orientations. To check for this, we run model (7) using the following dependent variables: a dummy variable for graduated mayors; a dummy for female mayors; the variable *age* (i.e., the age of the mayor); and a dummy for center-left oriented mayors. As shown in Panel A of Table 3 (and visualized in Figure 2), 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 implement 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 not accompanied by any changes in tax powers comparable to the introduction of the ICI property tax at the 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 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 (*administrative index*, *administrative jobs*$[A]$, *administrative 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 Panels B and C of Table 3. The geographical units used are the Italian provinces.[33] 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. On the contrary, for political skills, the coefficients are in general *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, the opposite of the behavior that we observe at the municipal level. In conclusion, the results at the municipal level 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.

[33]The 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 the provincial level, calculating the average values of the variables used by province.

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 for services, the opportunity cost of becoming a politician, or the degree of political competition. To control for these stories, we add to model (7) some covariates which allow controlling 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 level.[34] The results obtained after controlling for all these covariates are reported in columns (1), (3) and (5) of Table 4 for a-type politicians, and in columns (7) and (9) of Table 4 for p-type politicians. As shown, our results remain virtually unchanged in terms of the sign, magnitude, and significance after the introduction of these additional covariates.[35]

Controlling for alternative mechanisms Another potential threat to the causal interpretation of our results is the simultaneity between the municipal reforms and other factors that might have changed their role after decentralization. In the even columns of Table 4, we report the results obtained running model (9), in which we interact *decentralization* with five alternative mechanisms. First, we control for the role of the political scandals at the beginning of the '90s. These scandals might 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 class.[36] More in detail, we control for the 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 parties.[37] Turnout was as high as 77% and the proposal was approved in all Italian provinces, but with some variability. We define a variable measuring the share of voters in favor of abolishing public funding to parties ($\%\_referendum$) at the provincial level and interact this variable with *decentralization*. We interpret $\%\_referendum$ as a proxy for the willingness to change the political system in the different parts of the countries. Second, we control for the different endowment of social capital across Italian cities, which may influence voters' preferences. We consider the interaction between the number of employees in non-profit organizations per capita ($\#\_no\_profit$, a commonly used proxy for social capital, see Guiso et al., 2016) and *decentralization*.

Third, while our model emphasizes the demand effects on local politicians, there may also

[34] Galasso 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.

[35] In Table 4, we do not control for variables that capture the affiliation of mayors to the new parties that emerged after the political turmoil of the '90s. This is because these variables are potential dependent variables in our empirical model, and thus they should be treated as *bad controls*. However, as a robustness check, we rerun the regressions in Tables 4 adding dummy variables equal to one for those mayors affiliated to the new political parties that emerged during the years studied. As shown in Panel A of Table B.2, the results described in Table 4 remain entirely unchanged.

[36] The 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 epicenter in Milano, one of the richest cities in our sample.

[37] The referenda were eight, ranging from partially allowing the use of drugs to reducing the role of some ministries such as Health, Agriculture and Tourism. We pick the one directly related to political parties.

be supply-side effects driving, given that poor cities may have fewer candidates of $a-$type available. For this reason, we consider two additional interactions with *decentralization*: the number of firms per capita at the provincial level to capture the availability of entrepreneurs, and the number of graduate individuals out of the adult population at the provincial level to capture the availability of professionals. Fourth, 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 *decentralization* and the Herfindahl index of political competition at the local level. Finally, we control that our results are not driven by the North-South divide[38], adding to the model an interaction term between *decentralization* and a dummy variable equal to 1 for municipalities in center-north regions (i.e. Emilia Romagna, Liguria, Lombardia, Piemonte, Veneto, Lazio, Marche, Toscana, Umbria). Columns (2), (4), (6), (8), and (10) of Table 4 report the results of these robustness checks. As we can see, after adding these interaction terms, the coefficient for the interaction term between *decentralization* and fiscal autonomy is still significant, maintains the expected sign, and strongly increases in magnitude. Most of the coefficients of the other interaction terms are not statistically significant, a result that suggests that the alternative mechanisms considered cannot explain the main results of our paper.[39]

5.3 Additional robustness checks

Besides the main robustness checks described above, we discuss four other potential threats to our results. First, Panel B of Table B.2 shows that our results are unchanged if we exclude the regional capitals from our regressions. This robustness check enables us to control for the potential role of political campaign costs (e.g., Poutvaara and Takalo, 2007; Kendall et al., 2015), which may 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. Second, Panels A and B of Table B.4 demonstrate that the results are similar if we use measures of fiscal autonomy alternative to the pre-reform GDP. More specifically, in Panel A, we use the ratio between municipal own revenues (in terms of taxes and fees) and total revenues (including grants from Central government and fees). We average this ratio over the years 2000-2012, for which we have data, and use this average as a new definition of $fiscal\ autonomy$. In Panel B, we use as an alternative measure of $fiscal\ autonomy$ the revenues per capita from an experimental version of the property tax (Imposta Straordinaria sugli Immobili, ISI[40]), that was introduced

[38]Southern regions are on average poorer, 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.

[39]As explained in subsection 5.1.2, in this context, education does not seem to be a good proxy for administrative skills. The negative coefficient in front of the interaction term between *decentralization* and the share of college graduates in column 2 of Table 4 seems to further confirm this intuition. Besides, note that some of the control variables in Table 4 are time-varying, and thus they could potentially be affected by our treatment. We use time-varying variables because the presence of municipal fixed effects does not allow differently. As a robustness check, in columns (2), (4), (6), (8), and (10) of Table B.3, we repeat the same exercises using interaction terms between *decentralization* and time-invariant versions of the same control variables. As we can see, the results do not change.

[40]ISI was a smaller and experimental version of ICI, collected by the central government, and introduced one year before ICI. Data on ISI comes from an estimation, given that information on the revenues collected by the central government from every municipality is not available. In practice, the ISI revenues for every municipality have been estimated combining the total value of ISI revenues raised by the Italian government

in 1992 for one year only.[41] Third, in Panel C of Table B.4, we show that our main findings are confirmed even if we drop non-electoral years and we re-estimate equation (7) by using only data for electoral years.[42] Finally, Panel A of Table B.5 shows that the results are unchanged is we use alternative definitions of the dependent variables Y^a and Y^p which are mutually exclusive.[43]

6 Voters' welfare

6.1 Proxies for welfare

Our model predicts that rich communities will elect candidates with administrative skills and welfare will improve; while poor communities will elect candidates with political skills, with uncertain effects on welfare. In this section, we provide evidence on the ex-post performance of mayors, looking for proxies of welfare related to administrative/managerial skills. Specifically, we proxy Y^w by selecting two measures, one on the revenue side and the other on the expenditure side, that are independent of the number 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 $[A]$, 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). 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*), an indicator provided by Legambiente, an Italian independent environmental organization. Managing waste is an important task assigned to municipalities, on which mayors are easily

in 1992 with the less recent cadastral values available, which are those for 2008. Besides, note that the fact that ISI was introduced one year before ICI is not an issue for our strategy. First, ISI was collected by the central government, and it did not have any impact on municipal budget and fiscal autonomy. Second, we use the first year an election took place after the introduction of ICI as the first year of treatment. Given that ISI was introduced in 1992 and ICI replaced it in 1993, changing the first year of treatment with the first year an election took place after the introduction of ISI would not make any difference.

[41] As shown in Table A.5, these two measures for fiscal autonomy are highly correlated with GDP. This evidence suggests that given the limitations in the two measures for fiscal autonomy that are available, using GDP as a proxy for fiscal autonomy is reasonable.

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

[43] In the model, the a-type and the p-type are two mutually exclusive types of politicians, while in the empirical analysis we "mixed up" the two types. For example, a mayor from a profession with 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 apply the following four mutually exclusive definition of Y^a and Y^p: 1) $Y^a[A] = 1$ if *administrative jobs*$[A] = 1$ and *years political experience* $= 0$; 2) $Y^a[B] = 1$ if *administrative jobs*$[B] = 1$ and *years political experience* $= 0$; 3) $Y^p[A] = 1$ if *administrative jobs*$[A] = 0$ and *years political experience* > 0; 4) $Y^p[B] = 1$ if *administrative jobs*$[B] = 0$ and *years political experience* > 0.

evaluated by citizens and for which they are considered accountable.[44] 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.

We also account for the potential role of term limits. Specifically, 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 $decentralization \times fiscal\ autonomy$. This allows us to assess whether the effect of tax decentralization on the performance of mayors is due to re-election incentives or a selection effect. While the behavior of a mayor during the first term in office (when she can be re-elected) can be driven by both personal characteristics and by re-election incentives, in the second term in office the performance should be driven solely by personal characteristics as re-election incentives are no longer in place.

6.2 Results

The results of this exercise are reported in Table 5: 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).* 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.

The results in Table 5 suggest that tax decentralization reforms have differentially affected performance across rich and poor municipalities through the selection of different types of politicians. In fact, following the argument proposed by Gagliarducci and Nannicini (2013), when a mayor is not term-limited, the tax decentralization reforms can affect the performance of the municipal government both through electoral incentives and the selection of different types of mayors. On the other hand, when the mayor cannot be re-elected, the tax decentralization reforms can affect the performance only through the selection of different types of mayors. The results of Table 5 indicate that, when both the re-election and the selection mechanisms are in place (i.e., for first-term mayors), we do not find differences in the effect of tax decentralization reforms 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. Finally, as suggested by Gagliarducci and Nannicini (2013), calculating the difference between the coefficients of first-term and second-term mayors, it is possible to isolate the effect of tax decentralization on performance that works only through re-election incentives. The negative coefficients obtained calculating

[44] See, for instance, the echo in international media caused by the rubbish crisis in Naples in 2008, in Palermo in 2012, and in Rome in more recent years.

this difference suggests that re-election incentives cancel out the positive effect induced by the selection of different types of mayors, and induce the mayors to decrease both the speed of revenues collection and the amount of separate waste collection. This result is consistent with the literature that shows that electoral incentives induce politicians to cut taxes (Alesina and Paradisi, 2017) and reduce the number of resources destined to environmental policies (List and Sturm, 2010).

7 Concluding remarks

The literature on fiscal federalism supports the view that decentralization requires, to work properly, enough autonomous resources to finance local spending. The common explanation for this positive relationship 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 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 provide supporting evidence for this argument in the case of the Italian decentralization reforms of the '90s involving municipalities. The analysis shows 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 various alternative potential explanations. Interestingly, 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 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 for 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 to guarantee that central money keeps flowing to the community. Political skills and political connections are more important. In principle, these incentives could be counteracted by rule-based formulas that eliminate 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 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. Hence, the analysis in our paper suggests that decentralization may work well only if applied to subnational governments that have enough local resources. It would be interesting if future research could replicate our analysis to other countries affected by a decentralization process.

References

[1] Alesina A., Cassidy T. and Troiano U. (2018). "Old and young politicians." *Economica*.

[2] Ahmad E., Brosio G., eds. (2008). Handbook of Fiscal Federalism. Edward Elgar, London.

[3] Alesina A., Paradisi M. (2017). "Political Budget Cycles: Evidence from Italian Cities." *Economics&Politics*

[4] 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.

[5] 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.

[6] 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.

[7] Angrist J. D., Pischke J.-S. (2009). Mostly Harmless Econometrics. Princeton University Press.

[8] 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.

[9] Baldini G., Legnante G. (2000). Città al voto: i sindaci e elezioni comunali. Il Mulino, Bologna.

[10] Bardhan P., D. Mookherjee (2000). "Capture and Governance at Local and National Levels." *American Economic Review*, 90(2): 135-139.

[11] Bardhan P., D. Mookherjee, eds. (2006) Decentralization and Local Governance in Developing Countries: A Comparative Perspective, MIT Press.

[12] 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.

[13] Besley T. (2004). "Paying politicians: theory and evidence." *Journal of the European Economic Association* 2, 193-215.

[14] Besley T. (2005). "Political Selection." *Journal of Economic Perspectives*, 19(3): 43–60.

[15] Besley T. (2006). "Principled Agents? The Political Economy of Good Government". Oxford University Press, Oxford.

[16] Besley T., A. Case (1995). "Incumbent Behavior: Vote Seeking, Tax Setting and Yardstick Competition", *American Economic Review*, 85(1): 25-45.

[17] Besley, T. and Reynal-Querol M., (2011) "Do Democracies Select More Educated Leaders?," *American Political Science Review*, 105 (2011), 552–556."

[18] Besley, T., Montalvo J. G. and Reynal-Querol M., (2011). "Do Educated Leaders Matter?" *The Economic Journal.* Volume 121, Issue 554, pp. 205-227.

[19] Besley T., Smart M. (2007). "Fiscal restraints and voter welfare", *Journal of Public Economics*, 91(3-4): 755-773.

[20] Bettin Lattes G., Magnier A. (1995). "I nuovi sindaci: come cambia una carriera politica". Rivista Italiana di Scienza Politica, 25(1):91-118.

[21] Borck, R., Owings, S., (2003). "The political economy of intergovernmental grants." *Regional Science and Urban Economics* 33, 139–156.

[22] 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.*

[23] 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.

[24] Bordignon M., Colombo L., Galmarini U. (2008). "Fiscal federalism and lobbying". *Journal of Public Economics*, 92(12): 2288-2301.

[25] Bordignon M., Nannicini T., Tabellini G. (2016). "Moderating political extremism", *American Economic Review*, 106 (8), 2349-70.

[26] Bordignon M., Turati G. (2009). "Bailing out expectations and public health expenditure", *Journal of Health Economics*, 28 (2), 305-21.

[27] Bracco E., Lockwood B., Porcelli F. and Redoano M.(2015). "Intergovernmental grants as signals and the alignment effect: Theory and evidence" *Journal of Public Economics* 123 (2015) 78–91.

[28] 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.

[29] Brollo F., Nannicini T., Perotti R., Tabellini G. (2013). "The Political Resource Curse". *American Economic Review*, 103 (2013), 1759-1796.

[30] Brollo F. and Troiano U., (2016). "What Happens When a Woman Wins an Election? Evidence from Close Races in Brazil." *Journal of Development Economics.*

[31] Cai H. and Treisman D. (2005). "Does Competition for Capital Discipline Governments? decentralization, Globalization and Public Policy." *American Economic Review*, 95(3): 817-830.

[32] 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.

[33] 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.

[34] 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.

[35] Caselli F. and Morelli M. (2004). "Bad politicians". Journal of Public Economics, 88, 759-782.

[36] Curto M., Solé Ollé A., and Sorribas P. (2018). "Does electoral competition curb party favoritism?". American Economic Journal: Applied Economics, 10(4), 378-407, 2018.

[37] 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.

[38] Dal Bo' E., Finan F., Folke O. and Rickne J. (2017). "Who becomes a politician?" *Quarterly Journal of Economics*.

[39] Dalle Nogare C. and Kauder B., (2017). "Term limits for mayors and intergovernmental grants: Evidence from Italian cities." *Regional Science and Urban Economics* 64 (2017) 1–11.

[40] Daniele G. (2017). "Strike one to educate one hundred: Organized crime, political selection and politicians' ability." *Journal of Economic Behavior & Organization (2017).*

[41] Daniele G. and Geys B. (2015). "Organised crime, institutions and political quality: empirical evidence from Italian municipalities." *Economic Journal*, 125: F233–F255.

[42] 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.

[43] 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.

[44] Eyraud L., Lusinyan L. (2013). "Vertical fiscal imbalances and fiscal performance in advanced economies". *Journal of Monetary Economics*, 60(5): 571-587.

[45] Ferraz C., Finan F. (2009). "Motivating Politicians: the impacts of monetary incentives on quality and performance." NBER Working Paper No. 14906.

[46] Fisman, R. and R. Gatti (2002), "decentralization and Corruption: Evidence from US Federal Transfer Programs", *Public Choice*, 113: 25-35.

[47] Folke, O., and J. Rickne (2016) "The Glass Ceiling in Politics Formalization and Empirical Tests," *Comparative Political Studies*, , 567–599.

[48] 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.

[49] Gagliarducci S., Nannicini T. (2013). "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection". *Journal of the European Economic Association*, 11: 369-398.

[50] Gagliarducci S., Nannicini T. and P. Naticchioni (2010). "Moonlighting Politicians" *Journal of Public Economics*, 94(9-10), 688-699.

[51] Gagliarducci S., Paserman D. (2012). "Gender Interactions within Hierarchies: Evidence from the Political Arena" *Review of Economic Studies*, 79(3), 1021-1052.

[52] Galasso V., Nannicini T. (2011). "Competing on good politicians." *American Political Science Review*, 105(1): 79-99.

[53] Gamalerio M. (2019). "Fiscal rules and the selection of politicians: evidence from Italian municipalities." Working paper.

[54] Gamalerio M. (2020). "Do national political parties matter? Evidence from Italian municipalities." Working paper.

[55] Galiani, S., P. Gertler and E. Schargrodsky (2008). "School decentralization: Helping the Good Get Better, but Leaving the Poor Behind." *Journal of Public Economics*, 92(10-11): 2106 -2120.

[56] Gehlbach S., Sonin K. and Zhuravskaya E., (2010). "Businessman candidates." *American Journal of Political Science*, Volume 54, Issue 3, Pages 718-736, July 2010.

[57] 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.

[58] Guiso L., Sapienza P. and Zingales L. (2016). "Long-term Persistence." *Journal of the European Economic Association*.

[59] Hayeon Joo H. and Lee J., (2018). "Encountering female politicians." *Journal of Economic Behavior & Organization*, Volume 151, July 2018, Pages 88-122.

[60] Hines J. R. and R. H. Thaler (1995). "The Flypaper Effect". *Journal of Economic Perspectives*, 9(4): 217-226.

[61] Imbens G. W. and J. M. Wooldridge (2009). "Recent developments in the econometrics of program evaluation". *Journal of Economic Literature*, 47(1): 5-86.

[62] List J. and Sturm D. M. (2010). "How Elections Matter: Theory and Evidence from Environmental Policy." *Quarterly Journal of Economics*, 121(4), November 2006, 1249-1281.

[63] 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.

[64] 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.

[65] Key, V. O., (1956). "American State Politics: An Introduction." New York: Alfred A. Knopf.

[66] 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.

[67] Levitt, D.S., and J.M. Snyder. (1995). "Political Parties and the Distribution of Federal Outlays." *American Journal of Political Science,* 39: 958–80.

[68] Lockwood B. (2015). "Fiscal decentralization: A Political Economy Perspective", in E. Ahmad and G. Brosio (eds.), The Handbook of Fiscal Federalism, Edward Elgar.

[69] Martinez-Bravo M., (2017). "The Local Political Economy Effects of School Construction in Indonesia" " *American Economic Journal: Applied Economics*, 2017, 9(2): 256–289.

[70] Mattozzi A. and Merlo A. (2008). "Political careers or career politicians?" *Journal of Public Economics*, 92(3-4): 597-608.

[71] Myerson, Roger B., (2006). "Federalism and incentives for success of democracy." *Quarterly Journal of Political Science* 1, 3–23.

[72] Pasquino G. (2006). I sistemi elettorali. Il Mulino, Bologna.

[73] Persson T., Tabellini G. (1994). "Does centralization increase the size of government?". *European Economic Review*, 38(3-4): 765-773.

[74] Persson T., Tabellini G. (2000). "Political Economics: Explaining Economic Policy". Cambridge, MA: MIT Press.

[75] Poutvaara P., Takalo T. (2007). "Candidate quality". *International Tax and Public Finance*, 14: 7-27.

[76] Putnam R. (1993) Making Democracy Work. Civic tradition in Modern Italy. Princenton University Press.

[77] 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.

[78] Rodden J. (2002). "The dilemma of fiscal federalism: grants and fiscal performance around the world." *Americal Journal of Political Science*, vol. 46, No.3, July 2002

[79] Rodden J. (2006). "The Political Economy of Federalism" in Weingast B. and D. Wittman (eds.), Oxford Handbook of Political Economy, Oxford University Press.

[80] Rodden, J. A., Eskeland G. S., and Jennie L. (2003). "Fiscal decentralization and the Challenge of Hard Budget Constraints." Cambridge, MA: MIT Press.

[81] Rogoff K., Siebert A. (1988). "Elections and Macroeconomic Policy Cycles". *Review of Economic Studies*, 55:1-16.

[82] 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.

[83] Shugart M. S., M. P. Wattenberg (eds.) (2001). "Mixed Member Electoral Systems: The Best of Both Worlds?", Oxford: Oxford University Press.

[84] Treisman D. (2007). "The Architecture of Government. Rethinking Political decentralization", Cambridge: Cambridge University Press.

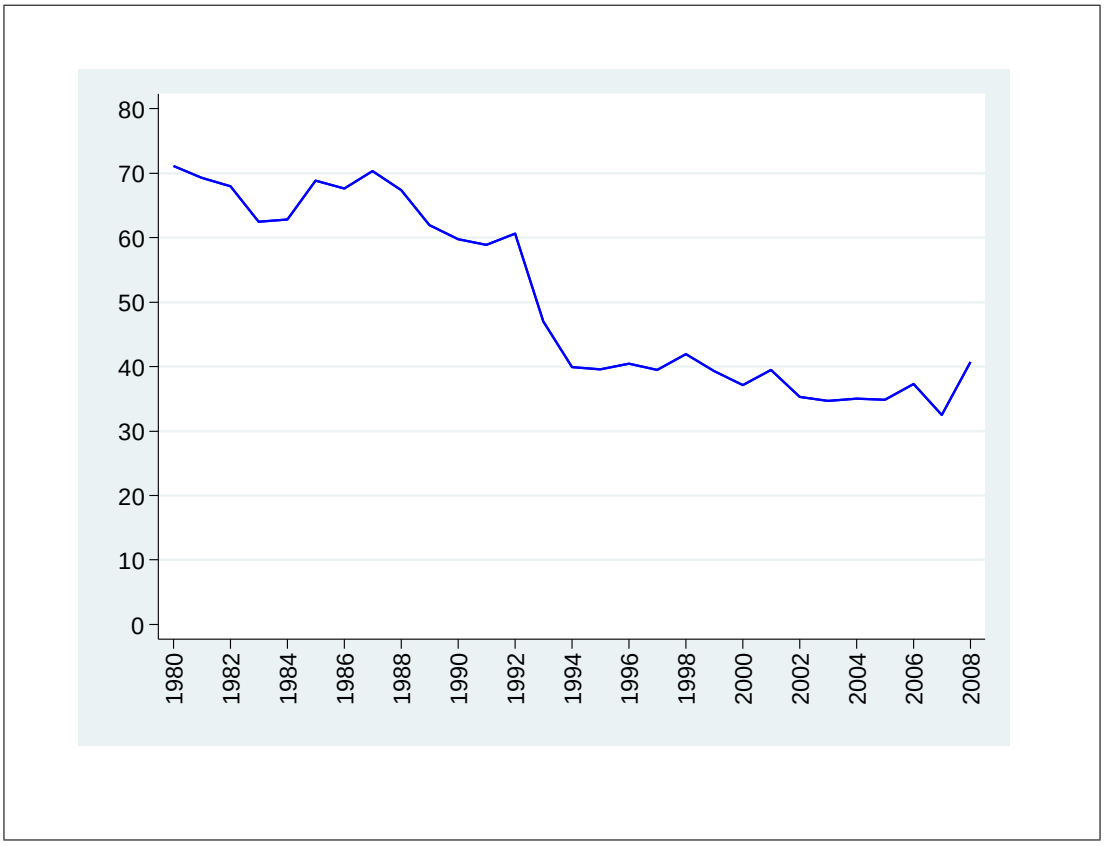
[85] Turati, G., Montolio, D., Piacenza, M. (2017). "Funding and School Accountability: The Importance of Private and Decentralised Public Funding for Pupil Attainment." *Fiscal Studies* 4; 2017 (38): 615-639.

[86] Weingast, B. R. (2009). "Second generation fiscal federalism: The implications of fiscal incentives" *Journal of Urban Economics* 65 (2009) 279–293.

[87] 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.

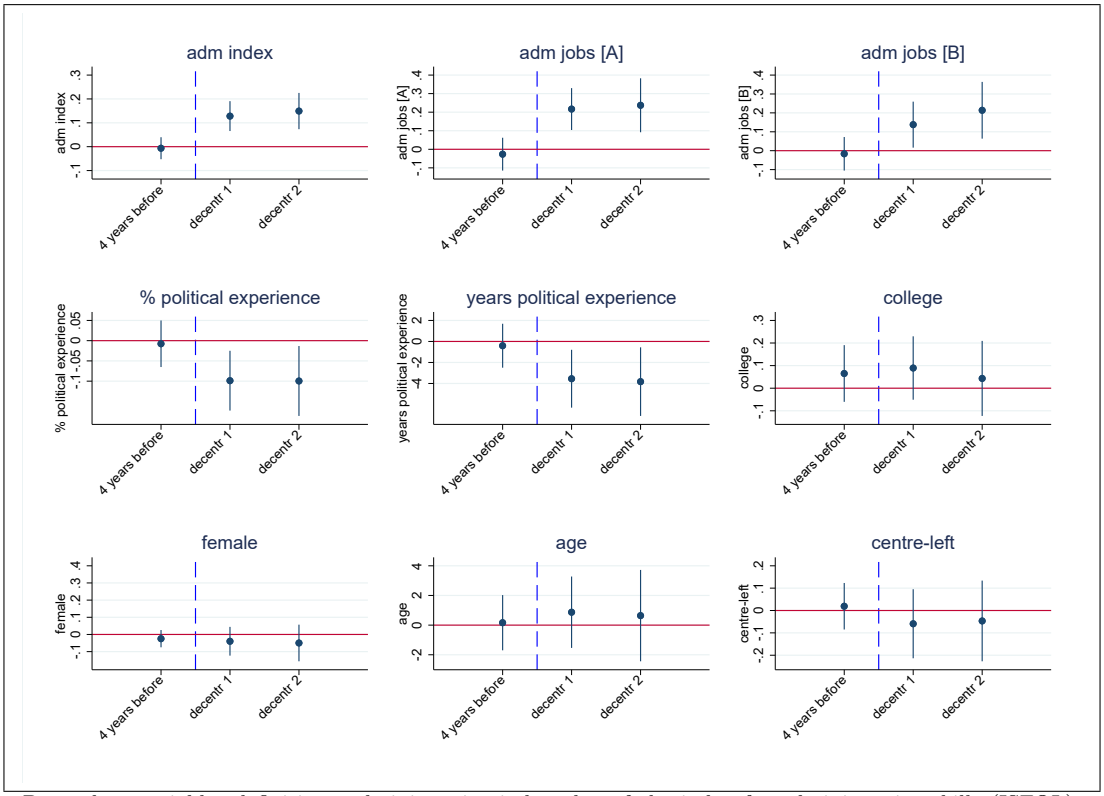
8 Tables and Figures

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



Source: ISTAT, Conto economico delle amministrazioni comunali.

Figure 2: Effect of decentralization on mayors' characteristics



Dependent variables definition: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % 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; 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. Independent variables definition: decentralization 1: dummy variable equal to 1 after the introduction of the property tax (ICI) interacted with gdp per capita measured in the years 1992-1993; decentralization 2: dummy variable equal to 1 after the introduction of the PIT surcharge interacted with gdp per capita measured in the years 1992-1993; all 4 years before: dummy variable equal to 1 for all 4 years before the introduction of the property tax (ICI) together interacted with gdp per capita measured in the years 1992-1993.

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

	(1)	(2)	(3)	(4)	(5)
	administrative index	administrative jobs [A]	administrative jobs [B]	% political experience	years political experience
Panel A: baseline model					
decentralization × fiscal autonomy	0.137***	0.246***	0.163***	-0.091***	-3.168***
	(0.039)	(0.052)	(0.049)	(0.022)	(0.856)
Observations	1,304	1,304	1,304	1,330	1,330
Panel B: two decentralization reforms					
decentralization 1 × fiscal autonomy	0.135***	0.244***	0.155***	-0.091***	-3.133***
	(0.039)	(0.052)	(0.049)	(0.022)	(0.841)
decentralization 2 × fiscal autonomy	0.155***	0.262***	0.230***	-0.092***	-3.436***
	(0.045)	(0.072)	(0.065)	(0.029)	(1.112)
Observations	1,304	1,304	1,304	1,330	1,330
Year FE	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes
Covariates	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: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % political experience = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years political experience: years of past political experience of the mayor at any institutional level. Independent variables definition: decentralization: dummy variable equal to 1 after the introduction of the property tax (ICI); fiscal autonomy: 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; decentralization 1: dummy variable equal to 1 after the introduction of the property tax (ICI); decentralization 2: dummy variable equal to 1 after the introduction of the PIT surcharge.

Table 2: Impact of decentralization on types of politicians (estimates of δ_1 and δ_2)
Two decentralization reforms and common trends

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	administrative index		administrative jobs [A]		administrative jobs [B]		% political experience		years political experience	
decentralization 1 × fiscal autonomy	0.128***	0.123***	0.217***	0.201**	0.138**	0.133*	-0.099***	-0.085*	-3.550**	-3.478*
	(0.032)	(0.044)	(0.058)	(0.080)	(0.062)	(0.079)	(0.038)	(0.048)	(1.405)	(1.897)
decentralization 2 × fiscal autonomy	0.149***	0.144***	0.237***	0.223**	0.214***	0.209**	-0.100**	-0.087	-3.827**	-3.758*
	(0.039)	(0.049)	(0.074)	(0.093)	(0.077)	(0.092)	(0.044)	(0.053)	(1.665)	(2.113)
all 4 years before × fiscal autonomy	-0.006		-0.026		-0.016		-0.008		-0.407	
	(0.024)		(0.045)		(0.045)		(0.029)		(1.067)	
1 year before × fiscal autonomy		-0.006		-0.037		-0.015		0.007		-0.392
		(0.040)		(0.076)		(0.076)		(0.041)		(1.619)
2 years before × fiscal autonomy		-0.025		-0.057		-0.036		0.008		-0.289
		(0.049)		(0.077)		(0.072)		(0.035)		(1.383)
3 years before × fiscal autonomy		-0.006		-0.032		-0.015		0.011		0.123
		(0.021)		(0.046)		(0.050)		(0.030)		(1.038)
4 years before × fiscal autonomy		-0.005		-0.019		-0.015		-0.020		-0.701
		(0.028)		(0.046)		(0.045)		(0.030)		(1.065)
Observations	1,304	1,304	1,304	1,304	1,304	1,304	1,330	1,330	1,330	1,330
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	No	No	No	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: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % political experience = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years political experience: years of past political experience of the mayor at any institutional level. Independent variables definition: decentralization 1: dummy variable equal to 1 after the introduction of the property tax (ICI); decentralization 2: dummy variable equal to 1 after the introduction of the PIT surcharge; fiscal autonomy: 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; all 4 years before: dummy variable equal to 1 for all 4 years before the introduction of the property tax (ICI) together; 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); 3 years before: dummy variable equal to 1 three years before the introduction of the property tax (ICI); 4 years before: dummy variable equal to 1 four years before the introduction of the property tax (ICI).

Table 3: The role of the electoral law
Placebo tests

	(1)	(2)	(3)	(4)
Panel A: other mayors' charcateristics				
	college	female	age	centre-left
decentralization × fiscal autonomy	0.018	-0.016	0.677	-0.077
	(0.051)	(0.029)	(0.927)	(0.065)
Observations	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
Panel B: administrative skills and education of Memebers of Parliament				
	administrative index	administrative jobs [A]	administrative jobs [B]	graduate
decentralization × fiscal autonomy	0.070	0.043	0.050	-0.054
	(0.046)	(0.069)	(0.059)	(0.064)
Observations	413	413	413	415
Year of election FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Covariates	No	No	No	No
Panel C: political skills of Memebers of Parliament				
	politician	# past terms	# past years	other offices
decentralization × fiscal autonomy	0.067	0.801***	3.085***	-0.031
	(0.042)	(0.224)	(0.839)	(0.066)
Observations	415	415	415	415
Year of election FE	Yes	Yes	Yes	Yes
Province 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 in Panel A. Provincial and year of election FE included in all models in Panel B and C. Dependent variables definition: Panel A: 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. Panel B: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; graduate=1 for graduate MPs. Panel C: 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. Independent variables definition: decentralization: dummy variable equal to 1 after the introduction of the property tax (ICI); fiscal autonomy: variable that captures the level of fiscal autonomy of the municipality in Panel A, averaged at provincial level in Panel B and C. In this case it is equal to gdp per capita measured in the years 1992-1993.

Table 4: Impact of decentralization on types of politicians (estimates of δ)
Adding control variables and interaction terms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	administrative index		administrative jobs [A]		administrative jobs [B]		% political experience		years political experience	
decentralization × fiscal autonomy	0.136***	0.312***	0.243***	0.651***	0.167***	0.389*	-0.092***	-0.283***	-3.093***	-8.173**
	(0.037)	(0.120)	(0.052)	(0.209)	(0.049)	(0.207)	(0.021)	(0.104)	(0.831)	(3.958)
decentralization × # firms		-0.019		-0.026		0.034		-0.018		-0.843
		(0.034)		(0.055)		(0.054)		(0.022)		(0.850)
decentralization × pol. comp.		-0.000*		-0.000**		-0.000		0.000**		0.009***
		(0.000)		(0.000)		(0.000)		(0.000)		(0.003)
decentralization × % college		-0.038***		-0.017		-0.032		-0.009		-0.094
		(0.014)		(0.025)		(0.029)		(0.013)		(0.415)
decentralization × # no profit		0.003		-0.012		0.006		0.008		0.296
		(0.008)		(0.016)		(0.016)		(0.008)		(0.244)
decentralization × % referendum		0.003		0.009		0.001		-0.004		-0.128
		(0.004)		(0.006)		(0.006)		(0.003)		(0.100)
decentralization × centre-north		-0.090		-0.065		-0.202		0.032		0.503
		(0.111)		(0.230)		(0.247)		(0.085)		(3.203)
Observations	1,304	1,304	1,304	1,304	1,304	1,304	1,330	1,330	1,330	1,330
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	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: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % political experience = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years political experience: years of past political experience of the mayor at any institutional level. Independent variables definition: decentralization: dummy variable equal to 1 after the introduction of the property tax (ICI); fiscal autonomy: 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 year and provincial level; pol. comp.: herfindal index of political competition measured at electoral mandate and municipal level; % college: percentage of municipal adult population with a college degree from 1991 Census; # no profit: number of employees in no profit organizations from 1991 Census; % referendum: % of yes at 1993 referendum on party funding; centre-north = 1 for municipalities in centre-north regions (i.e. Emilia Romagna, Liguria, Lombardia, Piemonte, Veneto, Lazio, Marche, Toscana, Umbria). Additional control variables added to the regression but not reported in the Table: in all columns: yearly municipal population not interacted with decentralization; share of yearly municipal population older than 65 years old not interacted with decentralization; # firms not interacted with decentralization, pol. comp. not interacted with decentralization. In columns (2), (4) and (6): yearly municipal population interacted with decentralization; share of yearly municipal population older than 65 years old interacted with decentralization. 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 5: Impact of decentralization on mayors' performance

	(1)	(2)
	revenues collection	waste collection
decentralization × fiscal autonomy, Term Limit=0	0.566	0.547
	(1.044)	(0.697)
decentralization × fiscal autonomy, Term Limit=1	11.491**	5.124**
	(4.715)	(2.191)
Difference	-10.925**	-4.577**
	(4.754)	(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: decentralization: dummy variable equal to 1 after the introduction of the property tax (ICI); fiscal autonomy: 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); yearly municipal population; share of yearly municipal population older than 65 years old; # firms: number of firms per capita at year and provincial level ; pol. comp.: herfindal index of political competition measured at electoral mandate and municipal level. All the variables have been interacted with Term Limit.

Appendix A: Descriptive statistics and correlations [For Online Publication]

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

Variable	Obs	Mean	Std. Dev.	Min	Max	Source
administrative index	1304	3.898	0.286	3.296	4.327	Italian Ministry of Domestic Affairs,
administrative jobs[A]	1304	0.292	0.455	0	1	"Anagrafe degli amministratori locali"
administrative jobs[B]	1304	0.212	0.409	0	1	
% political experience	1330	0.263	0.211	0	0.93	Italian Ministry of Domestic Affairs,
						"Anagrafe degli amministratori locali";
years political experience	1330	8.529	7.545	0	56	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,
						"Anagrafe degli amministratori locali"
female	1330	0.053	0.225	0	1	
age	1330	50.648	9.312	29	85	
centre-left	1335	0.53	0.49	0	1	Italian Ministry of Domestic Affairs,
						"Anagrafe degli amministratori locali";
						OPENOPOLIS; historical archive "Corriere
						della Sera" and "La Repubblica"
waste collection	796	11.082	9.856	0	52.1	Legambiente
revenues collection	817	67.83	13.53	13.6	91.4	Gagliarducci and Nannicini (2013)
GDPpc	1335	16298	3731	8851	26041	Istituto Tagliacarne
population	1335	171622	338180	15008	2800000	ISTAT
% over 65	1335	17.994	4.022	7.690	27.88	
# no profit	1335	8.382	3.688	1.349	17.514	
% college	1335	5.974	1.885	1.464	10.468	
# firms	1335	7.741	1.787	4.050	13.149	Italian Chambers of Commerce
political competition	1335	1963.897	439.553	1045.534	3657.212	Italian Ministry of Domestic Affairs,
						"archivio storico delle elezioni"; ITANES
% referendum	1335	90.1	3.2	81.6	96.6	Italian Ministry of Domestic Affairs,
						"archivio storico delle elezioni"
term limit	1335	0.155	0.362	0	1	Italian Ministry of Domestic Affairs,
						"Anagrafe degli amministratori locali"

Table A.2: descriptive statistics
Gdp above vs below the median

	(1)	(2)	(3)	(4)	(5)
	gdp above median		gdp below median		
Variable	mean	# municipalities	mean	# municipalities	p-value
population	241,370	44	103,422	45	0.054
% elderly	20.330	44	15.708	45	0.000
# firms	8.226	44	7.266	45	0.000
% college	6.359	44	5.597	45	0.057
# no profit	10.057	44	6.744	45	0.000
pol. comp.	1953.318	44	1974.240	45	0.668
% referendum	92.365	44	87.920	45	0.000
centre-north	1	44	0.422	45	0.000

Notes. The Table reports the munipal mean of the independent variables used in the analysis, distinguishing between 2 groups of municipalities: 1) gdp above median = municipalities with a gdp per capita above the median; 2) gdp below median = municipalities with a gdp per capita below the median. Variables reported: average municipal population, average share of municipal population older than 65 years old, average number of firms at provincial level, percentage of adult municipal population with a college degree from 1991 Census, number of employees in no profit organizations from 1991 Census, average herfindal index of political competition, share of yes at 1993 referendum on party funding, dummy variable for municipalities in centre-north regions.

Table A.3: Correlation between administrative skills and personal characteristics of the mayors

	(1)	(2)	(3)	(4)	(5)
	administrative index	college	female	age	centre-left
administrative index	1.0000				
college	0.0663	1.0000			
female	0.0573	-0.0337	1.0000		
age	0.1415	0.1181	-0.0826	1.0000	
centre-left	-0.2594	-0.0883	-0.0510	-0.1572	1.0000

Notes. Variable definition: administrative index: log of the index for administrative skills (ISFOL); 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.

Table A.4: Correlation between political experience and grants from higher levels of government

	(1)	(2)	(3)	(4)
Dependent variable	(log) Transfers per capita			
% political experience	0.233**	0.181***		
	(0.098)	(0.070)		
years political experience			0.009***	0.004*
			(0.003)	(0.002)
Observations	971	971	971	971
Outcome mean	510.2	510.2	510.2	510.2
Year FE	No	Yes	No	Yes
Municipal FE	No	Yes	No	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in columns 2 and 4. Dependent variable: log of transfers per capita from higher levels of government. Independent variables definition: % political experience = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years political experience: years of past political experience of the mayor at any institutional level.

Table A.5: Correlation between different measures for fiscal autonomy

	(1)	(2)	(3)
	gdp per capita	fiscal autonomy 2000-2012	ISI revenues per capita
gdp per capita	1.000		
fiscal autonomy 2000-2012	0.694	1.000	
ISI revenues per capita	0.702	0.654	1.000

Notes. Variable definition: gdp per capita: gdp per capita measured in the years 1992-1993; fiscal autonomy 2000-2012: ratio between municipal own revenues (taxes+fees) divided by total municipal revenues (taxes+fees+grants from higher levels of government) averaged over the period 2000-2012; ISI revenues per capita: per capita revenues from ISI (Imposta Straordinaria sugli Immobili), measured in hundred of euros.

Appendix B: Robustness checks [For Online Publication]

Table B.1: Impact of decentralization on types of politicians (estimates of δ_1 and δ_2)
Municipalities that change at least one mayor before reform

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Administrative and Political Skills									
	administrative index		administrative jobs [A]		administrative jobs [B]		% political experience		years political experience	
decentralization 1 × fiscal autonomy	0.115***	0.111**	0.181***	0.159**	0.101	0.087	-0.102***	-0.096*	-3.682**	-3.858*
	(0.035)	(0.045)	(0.056)	(0.074)	(0.062)	(0.073)	(0.039)	(0.050)	(1.458)	(2.003)
decentralization 2 × fiscal autonomy	0.134***	0.130**	0.187**	0.167*	0.170**	0.157*	-0.105**	-0.099*	-4.012**	-4.171*
	(0.042)	(0.052)	(0.075)	(0.090)	(0.078)	(0.089)	(0.048)	(0.057)	(1.778)	(2.261)
all 4 years before × fiscal autonomy	-0.004		-0.036		-0.027		-0.017		-0.633	
	(0.027)		(0.045)		(0.047)		(0.030)		(1.117)	
1 year before × fiscal autonomy		-0.000		-0.055		-0.038		-0.013		-0.983
		(0.039)		(0.071)		(0.070)		(0.043)		(1.728)
2 years before × fiscal autonomy		-0.023		-0.075		-0.056		-0.007		-0.691
		(0.056)		(0.075)		(0.073)		(0.037)		(1.446)
3 years before × fiscal autonomy		-0.002		-0.042		-0.026		0.001		-0.108
		(0.022)		(0.045)		(0.051)		(0.031)		(1.133)
4 years before × fiscal autonomy		-0.002		-0.027		-0.023		-0.028		-0.883
		(0.032)		(0.047)		(0.047)		(0.030)		(1.086)
Observations	1,169	1,169	1,169	1,169	1,169	1,169	1,195	1,195	1,195	1,195
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	No	No	No	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: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % political experience = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years political experience: years of past political experience of the mayor at any institutional level. Independent variables definition: decentralization 1: dummy variable equal to 1 after the introduction of the property tax (ICI); decentralization 2: dummy variable equal to 1 after the introduction of the PIT surcharge; fiscal autonomy: 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; all 4 years before: dummy variable equal to 1 for all 4 years before the introduction of the property tax (ICI) together; 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); 3 years before: dummy variable equal to 1 three years before the introduction of the property tax (ICI); 4 years before: dummy variable equal to 1 four years before the introduction of the property tax (ICI).

Table B.2: Impact of decentralization on types of politicians (estimates of δ)
Role of new political parties and electoral campaigns

	(1)	(2)	(3)	(4)	(5)
	administrative index	administrative jobs [A]	administrative jobs [B]	% political experience	years political experience
Panel A: Controlling for new political parties					
decentralization × fiscal autonomy	0.128***	0.224***	0.161***	-0.083***	-2.812***
	(0.038)	(0.054)	(0.049)	(0.019)	(0.759)
Forza Italia	0.162***	0.249**	0.132	-0.118***	-3.308***
	(0.046)	(0.110)	(0.118)	(0.040)	(1.268)
Lega Nord	0.105**	0.277*	0.081	-0.172***	-6.498***
	(0.053)	(0.144)	(0.136)	(0.048)	(1.621)
Observations	1304	1304	1304	1330	1330
Panel B: Role of electoral campaigns (excluding regional capitals)					
decentralization × fiscal autonomy	0.142***	0.253***	0.253***	-0.089***	-2.724***
	(0.034)	(0.049)	(0.049)	(0.023)	(0.908)
Observations	1,081	1,081	1,081	1,105	1,105
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: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % political experience = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years political experience: years of past political experience of the mayor at any institutional level. Independent variables definition: decentralization: dummy variable equal to 1 after the introduction of the property tax (ICI); fiscal autonomy: 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: yearly municipal population; share of yearly municipal population older than 65 years old; # firms at year and provincial level; herfindal index of political competition measured at electoral mandate and municipal level.

Table B.3: Impact of decentralization on types of politicians (estimates of δ)
Adding time invariant control variables and interaction terms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	administrative index		administrative jobs [A]		administrative jobs [B]		% political experience		years political experience	
decentralization × fiscal autonomy	0.137***	0.295**	0.246***	0.599***	0.163***	0.413**	-0.091***	-0.321***	-3.168***	-9.956*
	(0.039)	(0.118)	(0.052)	(0.212)	(0.049)	(0.206)	(0.022)	(0.116)	(0.856)	(5.085)
decentralization × # firms		-0.012		-0.003		0.019		0.021		0.821
		(0.023)		(0.034)		(0.033)		(0.017)		(0.806)
decentralization × pol. comp.		-0.000**		-0.000		-0.000		0.000		0.003
		(0.000)		(0.000)		(0.000)		(0.000)		(0.003)
decentralization × % college		-0.035***		-0.009		-0.026		-0.008		-0.059
		(0.014)		(0.027)		(0.031)		(0.013)		(0.452)
decentralization × # no profit		0.002		-0.007		0.006		0.007		0.300
		(0.008)		(0.018)		(0.016)		(0.008)		(0.250)
decentralization × % referendum		0.003		0.001		-0.000		-0.002		-0.088
		(0.003)		(0.006)		(0.006)		(0.002)		(0.088)
decentralization × centre-north		-0.091		-0.049		-0.187		-0.012		-1.270
		(0.107)		(0.225)		(0.236)		(0.093)		(3.841)
Observations	1,304	1,304	1,304	1,304	1,304	1,304	1,330	1,330	1,330	1,330
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Dependent variables definition: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % political experience = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years political experience: years of past political experience of the mayor at any institutional level. Independent variables definition: decentralization: dummy variable equal to 1 after the introduction of the property tax (ICI); fiscal autonomy: 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 municipal level from the 1991 Census; pol. comp.: herfindal index of political competition at municipal level, average over the pre-treatment period; % college: percentage of municipal adult population with a college degree from the 1991 Census; # no profit: number of employees in no profit associations at the municipal level from the 1991 Census; % referendum: % of yes at 1993 referendum on party funding; centre-north = 1 for municipalities in centre-north regions (i.e. Emilia Romagna, Liguria, Lombardia, Piemonte, Veneto, Lazio, Marche, Toscana, Umbria). Additional control variables added to the regression but not reported in columns (2), (4) and (6): municipal population from the 1991 Census interacted with decentralization; share of municipal population older than 65 years old from the 1991 Census interacted with decentralization. All control variables enter only in interaction terms because they are time invariant, so they are already captured by municipal fixed effects.

Table B.4: Impact of decentralization on types of politicians (estimates of δ)
Alternative measure for fiscal autonomy and term level data

	(1)	(2)	(3)	(4)	(5)
	administrative index	administrative jobs [A]	administrative jobs [B]	% political experience	years political experience
Panel A: fiscal autonomy = ratio municipal revenues-total revenues (2000-2012)					
decentralization × fiscal autonomy	0.305***	0.586***	0.416***	-0.209***	-6.894***
	(0.085)	(0.106)	(0.109)	(0.047)	(1.608)
Observations	1,304	1,304	1,304	1,330	1,330
Panel B: fiscal autonomy = ISI revenues per capita					
decentralization × fiscal autonomy	0.217***	0.447***	0.267***	-0.153***	-5.412***
	(0.066)	(0.103)	(0.103)	(0.056)	(1.791)
Observations	1,235	1,235	1,235	1,257	1,257
Year FE	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes
Panel C: Term level data					
decentralization × fiscal autonomy	0.212***	0.270**	0.223**	-0.090**	-3.132*
	(0.056)	(0.113)	(0.089)	(0.043)	(1.736)
Observations	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 in Panels A and B, municipal and electoral FE in all models in Panel C. Dependent variables definition: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; % political experience = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years political experience: years of past political experience of the mayor at any institutional level. Independent variables definition: decentralization: dummy variable equal to 1 after the introduction of the property tax (ICI); fiscal autonomy in Panel A: ratio between municipal own revenues (taxes+fees) divided by total municipal revenues (taxes+fees+grants from higher levels of government) averaged over the period 2000-2012; fiscal autonomy in Panel B: per capita revenues from ISI (Imposta Straordinaria sugli Immobili), measured in hundred of euros; fiscal autonomy in Panel C: gdp per capita measured in the years 1992-1993. Additional control variables added to the regression but not reported in the Table: yearly municipal population; share of yearly municipal population older than 65 years old; # firms at year and province level, herfindal index of political competition measured at electoral mandate and municipal level.

Table B.5: Alternative definitions dependent variables, and impact of decentralization on alignment and independent mayors

	(1)	(2)	(3)	(4)
Panel A: Alternative definitions dependent variables				
	Y^a [A]	Y^p [A]	Y^a [B]	Y^p [B]
decentralization × fiscal autonomy	0.195***	-0.301***	0.113***	-0.308***
	(0.040)	(0.042)	(0.035)	(0.048)
Observations	1,304	1,304	1,304	1,304
Panel B: Alignment and independent mayors				
	Aligned mayor		Civic list	
decentralization × fiscal autonomy	-0.127	-0.136	0.072***	0.079**
	(0.087)	(0.090)	(0.028)	(0.032)
Observations	1,330	1,330	1,330	1,330
Year FE	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes

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: Y^a [A] = 1 if administrative jobs [A] = 1 and years political experience = 0; Y^a [B] = 1 if administrative jobs [B] = 1 and years political experience = 0; Y^p [A] = 1 if administrative jobs [A] = 0 and years political experience > 0; Y^p [B] = 1 if administrative jobs [B] = 0 and years political experience > 0; administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants; years political experience: years of past political experience of the mayor at any institutional level. Panel B: aligned mayor =1 if mayor aligned with national and/or regional government; civic list = 1 if mayor comes from a Civic List independent from national political parties. Independent variables definition: decentralization: dummy variable equal to 1 after the introduction of the property tax (ICI); fiscal autonomy: 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: yearly municipal population; share of yearly municipal population older than 65 years old; # firms at year and province level, herfindal index of political competition measured at electoral mandate and municipal level.

Table B.6: Impact of decentralization on types of politicians (estimates of δ_1 and δ_2)
Two decentralization reforms, change the mayor vs did not

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Municipalities	All	Change mayor	No change mayor	All	Change mayor	No change mayor	All	Change mayor	No change mayor
					Panel A: Administrative Skills				
Dependent variable	administrative index			administrative jobs [A]			administrative jobs [B]		
decentralization 1 × fiscal autonomy	0.135***	0.094***	0.190***	0.244***	0.246***	0.241***	0.155***	0.123**	0.194**
	(0.039)	(0.035)	(0.066)	(0.052)	(0.064)	(0.085)	(0.049)	(0.058)	(0.083)
decentralization 2 × fiscal autonomy	0.155***	0.134**	0.175**	0.262***	0.306***	0.217**	0.230***	0.240***	0.226**
	(0.045)	(0.057)	(0.069)	(0.072)	(0.098)	(0.102)	(0.065)	(0.088)	(0.095)
Observations	1,304	719	585	1,304	719	585	1,304	719	585
p-value difference	0.431	0.409	0.529	0.707	0.483	0.512	0.057	0.077	0.425
					Panel B: Political Skills				
	% political experience			years political experience					
decentralization 1 × fiscal autonomy	-0.091***	-0.107***	-0.071**	-3.133***	-3.940***	-2.204**			
	(0.022)	(0.030)	(0.036)	(0.841)	(1.263)	(1.100)			
decentralization 2 × fiscal autonomy	-0.092***	-0.149***	-0.038	-3.436***	-5.946***	-0.877			
	(0.029)	(0.036)	(0.046)	(1.112)	(1.442)	(1.582)			
Observations	1,330	731	599	1,330	731	599			
p-value difference	0.960	0.245	0.282	0.689	0.070	0.207			
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipal FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	No	No	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: Panel A: administrative index: log of the index for administrative skills (ISFOL); administrative jobs [A], jobs included: school managers, entrepreneurs, directors, engineers and architects, veterinarian, dentists, traders, biologists, university professors, surveyor, business consultants and pharmacists; administrative jobs [B], jobs included: school managers, entrepreneurs, directors, engineers and architects, traders, surveyor and business consultants. Panel B: % political experience = years of past political experience at any institutional level as a fraction of working age of the mayor (i.e. age-17 years); years political experience: years of past political experience of the mayor at any institutional level. Independent variables definition: decentralization 1: dummy variable equal to 1 after the introduction of the property tax (ICI); decentralization 2: dummy variable equal to 1 after the introduction of the PIT surcharge; fiscal autonomy: 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. Municipalities kept in the regressions: All = all municipalities; Change mayor = municipalities that change mayor when the PIT surcharge is introduced; No change mayor = municipalities that do not change mayor when the PIT surcharge is introduced.

Appendix C: Proofs theoretical model [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: $\underline{y}^j = \tau(\alpha)(-\frac{1}{2\psi} + \overline{\delta}^j)$; $\overline{y}^j = \tau(\alpha)(\frac{1}{2\psi} + \overline{\delta}^j)$, $\underline{x}^j = \alpha t(-\frac{1}{2\psi} + \overline{\theta}^j)$, $\overline{x}^j = \alpha t(\frac{1}{2\psi} + \overline{\theta}^j)$. Let $k = \frac{\psi^2}{\alpha t \tau(\alpha)}$. By the convolution theorem, under A.1, $f(z^j)$ is:

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

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

8.1 Derivation of r_1^{j*}.

Rewriting Eq.(4) fully,

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

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

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

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

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

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 $E'(R^j) = 0$,

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

that solving for r_1^j, gives

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

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 $E(R^{j^*}) \geq \overline{R}$ 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:

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

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

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

Notice from (5), that (**11**) also implies $r_1^{k*} > 0$. Hence, as stated in Proposition 1, for $R' > \overline{R}$ 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 $r_1^{j*} = 1 - \frac{\lambda \overline{R}\psi}{\tau}\left[\frac{E(z^o)}{E(z^j)}\right]$.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

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

Solving:

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

which is positive if

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

This holds if

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

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

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

Consider now

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

which is positive if

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

This holds if

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

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

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

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

- For p-type:

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

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

and

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

Finally:

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

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

QED.

Proof of Proposition 3

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

$$U^a=\frac{\lambda\overline{R}\psi}{\tau}E(z^o)-\lambda\overline{R}+\lambda(\pi(\frac{1}{2}(E(z^a)+E(z^a|z^a\geq E(z^a))+ \tag{12}$$
$$+\lambda(1-\pi)((\frac{1}{2}+\frac{\psi}{\tau}h)E(z^a|z^a\geq E(z^p))+(\frac{1}{2}-\frac{\psi}{\tau}h)E(z^p))$$

$$U^p=\frac{\lambda\overline{R}\psi}{\tau}E(z^o)-\lambda\overline{R}+\lambda((1-\pi)(\frac{1}{2}(E(z^p)+ \tag{13}$$
$$+E(z^p|z^p\geq E(z^p))+\lambda\pi((\frac{1}{2}-\frac{\psi}{\tau}h)E(z^p|z^p\geq E(z^a))+$$
$$+(\frac{1}{2}+\frac{\psi}{\tau}h)E(z^a))$$

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:

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

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

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

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

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 $\overline{\theta}^{a*} > \overline{\theta}^{p*} > 0$ such that $\pi \to 0$ and $\overline{\theta}^p \leq \overline{\theta}^{p*}$ implies $\partial U/\partial t < 0$ and $\pi \to 1$ and $\overline{\theta}^a \geq \overline{\theta}^{a*}$ implies $\partial U/\partial t > 0$. This proves (i). Differentiating $\partial U/\partial t$ with respect to π :

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

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:

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

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

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

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

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

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

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

where $A = E(R^a)$, $B = E(R^p)$, $E(z^a) = a$, $E(z^p) = p$; $k = \frac{\lambda\overline{R}\psi}{\tau}$. 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:

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

which can be rewritten as:

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

Let $A + B = x$; $\lambda\overline{R}$ $+(a+p)(1-k) = b; 2(a-p)^2k(1+k) = c$, 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 $k \to 1$, $Q \to \frac{\tau}{4\psi}$. 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 $a \to p$. Thus, the only economic significant solution is represented by the positive root. Under the positive root, $A + P = \frac{n(b+\sqrt{b^2-4c}))}{2\overline{w}(1+\lambda)}$. 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:

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

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

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

for $E(z^a) \geq (<)E(z^p)$. Notice that $(1 - \frac{\lambda\overline{R}\psi}{\tau^2}) = \frac{1}{\tau}(1 - \alpha t - \frac{\lambda\overline{R}\psi}{\tau})$ 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 negative provided that $\overline{\theta}^a - \overline{\delta}^a$ is not much larger than $\overline{\delta}^p - \overline{\theta}^p$. 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:

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

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