

Decentralization, Vertical Fiscal Imbalance, and Political Selection

Massimo Bordignon*

Department of Economics and Public Finance
Catholic University, Milan & CESifo
massimo.bordignon@unicatt.it

Matteo Gamalerio

Department of Economics
University of Warwick / Coventry / UK
m.gamalerio@warwick.ac.uk

Gilberto Turati

Department of Economics and Statistics
University of Turin & CIFREL, Catholic University, Milan
gilberto.turati@unito.it

March 2014

Abstract

In a career-concern model of politics with endogenous candidacy and different types of politicians, following a decentralization reform, politicians with different skills are elected in municipalities characterized by different levels of autonomous resources. As an effect, consumer welfare increases only, or mainly, in richer municipalities. We test these predictions by exploiting the differentiated reduction in Vertical Fiscal Imbalance in Italian municipalities, due to the strong difference in the tax base, following the decentralization reforms of the '90s. Results strongly support our predictions and are robust to several alternative stories.

Keywords: decentralization, vertical fiscal imbalance, quality of politicians

JEL codes: D72, D78

*Corresponding author: Department of Economics and Public Finance, Catholic University, Largo A. Gemelli 1, I-20123 Milano, Italy. Ph.: +39.02.7234.2694; Fax: +39.02.7234.2781; E-mail: massimo.bordignon@unicatt.it.

1 Introduction

As is well known, the case for fiscal decentralization is in general terms rather weak. The theoretical literature typically emphasizes the potentially large efficiency gains deriving from devolving to local governments the provision of important public services, in terms of better representation of local preferences and better accountability of politicians (e.g., Lockwood, 2008 and 2013, for recent surveys). The same ideas lie behind the wide support that fiscal decentralization usually receives in political circles, international organizations and development agencies, which in turn helps explaining the recent wave of decentralization in both developed and developing countries (e.g., Treisman, 2007; Bardhan and Mookerjee, 2006)¹. The empirical evidence is however 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, 2006). Understanding under which conditions fiscal decentralization is able to fulfill its promises represents therefore an important area for research.

On this matter, a paramount role seems to be played by the mismatch between own revenues and expenditure at the local level: higher levels of "vertical fiscal imbalances" (VFI) - i.e., higher shares of transfers in the local government budgets - are typically shown to be associated with poorer local governments' performance (e.g., Ahmad and Brosio, 2008). For instance, it has long been known that the propensity to spend out of transfers by local governments is higher, possibly leading to inefficient levels of public expenditure in communities largely financed with grants (the so-called "fly-paper effect"; e.g., Hines and Thaler, 1995, and Dahlberg et al., 2008). Similarly, a substantial empirical evidence shows that financial instability and soft budget constraints problems are more likely to occur when VFI is high than when it is low (e.g., Rodden et al., 2003, Eyraud and Lusinyan, 2011). Fisman and Gatti (2002) suggest that corruption is also higher when VFI is higher, and in different contexts such as Uganda and Argentina, Reinikka and Svensson (2004) and Galiani et al. (2008) show that, *ceteris paribus*, the quality of education provision at the local level is also poorer in localities where VFI is higher.

But *why* this is the case has never been made very clear in the literature. It is intuitive, for example, that local governments largely financed with transfers may not internalize entirely the cost of spending, presumably because citizens have less incentives in controlling how much money is spent in public services if a large part of this money does not come from their own pockets². But it is still unclear why citizens should tolerate a higher level of corruption, or a lower quality of services, just because the latter are largely financed with resources coming from outside the local community. Similarly, it is often argued that local politicians are more easily captured by local interests, leading to higher corruption or lower quality of services under decentralization, a point which goes back at least to the *Federalist*

¹Indeed, Treisman (2007: 3-4) estimates in several hundred million dollars the total sum that *each year* international organizations, banks, development agencies, single states etc. donate or lend to developing countries in order to support decentralization of fundamental services.

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

*papers*³. But, again, it is not obvious why this "capture" should be easier to occur under higher level of VFI.

A possible explanation, that we explore and emphasize in the present paper, is that there may be a relationship between the degree of VFI and the skills of (local) politicians. Our main idea, more formally discussed in Section 2, is the following. In a decentralized setting where most resources still comes from the center, the main task of a local politician lies in making sure that these resources keep flowing to the local community. This generally requires different political skills (say, strong party connections with the center, particular bargaining abilities, and extended family networks) than those of a good administrator of local matters. And given the choice, voters of communities with high degrees of VFI would rationally prefer the former type of politician to the latter. On the contrary, in communities where most of the resources are generated by the community itself (i.e., with a low degree of VFI), voters would rather prefer a good administrator than a local politician with strong connections. More generally, there can also be a self-selection effect. Anticipating voters' preferences, candidates of different abilities may decide to enter the local political arena in communities with different levels of VFI. As a result, as formally shown in Section 2, tax decentralization might have opposite effects in rich and poor communities. It increases voters' welfare in rich communities, as it also attracts politicians of higher administrative skills, while it reduces welfare in poorer communities as the quality of politicians does not change and their skills become less useful to voters. Interestingly, we prove that this occurs even with a "compensated" tax reform (as the Italian one discussed below) that leaves (at statutory tax rates) the total revenues of local governments unchanged, so that what basically changes with the reform is only the degree of VFI across municipalities.

We then take our theoretical results to the data, studying the Italian experience of decentralization in the '90s. This case study is particularly suitable to test our ideas. For, as explained in more detail in Section 3, the 1993 decentralization reform changed both the electoral system and the funding of municipalities, introducing a new tax source, the municipal property tax, that dramatically improved the financial prospects of municipalities. But while the electoral reform affected all municipalities in the same way, the tax reform had very different effects across the country. In the richest cities, because of their higher tax base, the newly introduced property tax made these municipalities almost entirely financially independent from the center; in the poorest ones, the effect was minimal and municipalities kept receiving most of their resources in the form of grants. On the basis of our theoretical argument, we expect that this should have affected both the selection of local politicians (inducing an abrupt change in the characteristics of the politicians in the richer communities and less or no effect in the poorer ones), and the quality of services (resulting in an improvement mainly, or only, in the richer municipalities).

To test our hypotheses, we then collect an extensive data set on the personal characteristics of the mayors of the main Italian cities both before and after the reform, as well as on other economic and political features of the municipalities⁴. Focusing only on the

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

⁴We consider all cities (89) that are also "capoluogo di provincia" (literally, chief provincial towns) in

mayor and not on the general characteristics of the local political class, aside from data availability, is justified by the paramount role in municipal policy attributed to mayors by the 1993 reform (again see Section 3 for detail). Departing from the political economy literature that usually proxies "quality" of politicians with either their level of education or their income (e.g., Gagliarducci and Nannicini, 2013; Galasso and Nannicini, 2011), we then construct several proxy measures for different "types" of politicians, distinguishing between mayors with "political skills" and mayors with "administrative skills". We also consider two separate indicators for the "ex-post" quality of urban policies, looking both at a specific and easily measurable output indicator that is related to one of the main services offered by Italian municipalities (separate waste collection), and at the probability of completing the term in office by the elected mayor, an indicator of efficiency for Italian municipalities that has been recently explored in the literature (e.g., Gagliarducci and Paserman, 2012).

Results are strongly supportive of our hypotheses. In municipalities where the electoral reform was accompanied by a large increase in autonomous resources, the ex-ante characteristics of politicians changed dramatically in the aftermath of the reform. A much larger percentage of elected mayors came from top administrative professions in the private sector. There is also some evidence that in these municipalities the higher quality of the local political class was also reflected in a higher quality of policies ex-post, and that this improvement was effectively due to a "selection effect" on local politicians and not to a stronger "disciplining effect" in richer communities⁵. On the contrary, we observe no or very little effects of the electoral reform in poorer communities, both in terms of the ex-ante skills of politicians and ex-post quality of policy. The causal impact of the funding reform is confirmed by placebo tests aimed at verifying the absence of any effects before it. Moreover, we also provide a number of robustness tests for these results, checking whether they were not driven by other factors that the literature typically associates with a better selection of politicians, or by other phenomena occurring in Italy in the same period. For instance, our results hold even controlling for the degree of competitiveness in the local electoral competition (e.g., Galasso and Nannicini, 2011), for the endowment of "social capital" at the municipal level (e.g., Guiso et al., 2011), for the higher costs of electoral campaigns in the richest cities, and for the changing political scenario in the mid '90s Italy, with the birth of new political parties. More importantly, they also hold in the case of another, smaller, reform in the financing of Italian municipalities (a municipal surcharge on the Personal Income Tax) that was introduced in 1999 at unchanged electoral rules (e.g., Bordignon and Piazza, 2010).

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

This study is linked to different lines of research. Beside the fiscal federalism literature,

the Italian ordinary statute regions.

⁵The terminology is borrowed from Besley and Smart (2007). We use an institutional feature of the Italian municipal electoral system, the presence of a term limit on mayors, to discriminate between these two effects.

our work is clearly related to the recent strand of research in political economics that focuses on the effects of political institutions on the selection of politicians (e.g., Besley, 2004, 2005, 2006; Caselli and Morelli, 2004; Poutvaara and Takalo, 2007; Mattozzi and Merlo, 2008; Gagliarducci et al., 2010). While most effort in this literature has been devoted so far to address the relationship between compensation and quality of politicians, the insight is clearly much more general and could be applied to other types of institutions, including decentralization. Finally, the idea that the features of the local political class may depend on the financial characteristics of the communities is probably not new, but to the best of our knowledge has not been formalized and explicitly tested before. The only exception is a recent work by Brollo et al. (2013) on Brazilian municipalities. Our work is related to theirs, but there are some important differences. First, they suggest that lower transfers always lead to a higher quality of politicians and higher voters' welfare, while in our model with different types of politicians this only occurs in rich communities. Second, they only study changes in transfers, while we consider 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.

The remainder of the paper is organized as follows. Section 2 presents a theoretical model of self-selection of local politicians under different local financial conditions that captures our main idea. Section 3 presents in more details the institutional characteristics of the Italian reforms. Section 4 discusses our econometric strategy and presents our data set, with some preliminary evidence. Section 5 is devoted to our main results, also discussing several robustness tests. Section 6 concludes.

2 The Model

The framework is a career-concern model of politics as in Brollo et al. (2013), but extended to different types of politicians with specialized skills. Thus, consider a 2 period economy, $t = 1, 2$, where t indexes the period, to which we add a self-selection stage later on. In this economy, at the beginning of the first period, an incumbent politician is in charge; at the end of this period an election takes place and either the incumbent or an opponent is elected to rule for the second period⁶. Politicians only care about collecting as much rents as possible from office. We let R_t indicate the rents appropriated by an incumbent in period t . Politicians come of two types, $j = a, p$; in a sense to be made more precise below, a -type is on average better in organizing local services (he has more "administrative" skills), p -type is on average better in raising money from the center (he has more "political" skills). Our basic point here is indeed that these are quite different skills, require a different type of backgrounds and specialization, and are therefore typically distributed differently across the population of (potential) politicians.

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

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_t . The utility of the voter over the two periods is then just:

$$U = g_1 + \lambda E(g_2) \quad (1)$$

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_t 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 use these resources; and negatively, on the rents that the politician in charge diverts to his advantage and to the detriment of voters. More specifically, in period 1, when an incumbent politician of type j is in charge, we assume:

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

where r_1^j is the rate of rents extraction in period 1 by a politician of type j , t is the exogenously given local tax rate ($0 < t < 1$) and α is the municipal tax base, with $\bar{\alpha} > \alpha > \underline{\alpha} > 0$. $\tau(\alpha) > 0$ is the transfer received by the center. Note that we write τ as a function of α , as grants to local governments in most countries have a redistributive component; they are larger in poorer communities ($\tau'(\alpha) < 0$)⁷.

Eq. (2) implies that g_1^j not only depends on the revenues accruing to the municipality, $\alpha t + \tau(\alpha)$, but also on the ability of politician j to use these resources, captured here by the couple (θ^j, δ^j) . More specifically, a given amount of local resources, αt , can generate a higher level of public good production if it is managed by a politician j that has a higher level of administrative skills than another politician k , $\theta^j > \theta^k$. This politician may also be better in managing the resources coming from the center in the form of grants. But we assume that on transfers political skills δ^j dominate, as a politician with higher political skills may be more able to get extra resources from the center, or to convince the center to directly finance some components of local expenditure⁸. Notice from Eq. (2) that we also assume that a politician can divert to his advantage the extra resources that he himself generates. This is intuitive: a politician with high political skills may cash some of the extra transfers he brings home; a politician with high technical skills (say, an architect) may divert some of the funds that he knows how to use better to his private associates, and so on. Finally, for analytical convenience, we do not allow politicians to take different rents from the different sources of financing; the same rate r_1^j applies to both sources.

Both θ^j and δ^j follow an independent uniform distribution function with density ψ and average $\bar{\theta}^j > 0$, $\bar{\delta}^j > 0$, respectively. In keeping with the discussion above, we assume $\bar{\theta}^a > \bar{\delta}^a$, $\bar{\delta}^p > \bar{\theta}^p$, $\bar{\theta}^a > \bar{\theta}^p$, $\bar{\delta}^a < \bar{\delta}^p$. This captures the idea that a -types are "better" on average in producing local services out of local resources, while p -types are "better" on

⁷For notational simplicity, in what follows we drop the dependence of τ on a when not needed for the argument.

⁸As an interpretation, one may think of $\tau(\alpha)$ as the transfer accruing to the local government as the result of the application of some rule-based grant formula, and to δ^j as a multiplicative component that depends on the political skills of the local politician.

average in raising extra resources from the center.

Let $z^j = \alpha t \theta^j + \tau(\alpha) \delta^j$ be the total municipal "productive" revenues generated by an incumbent of type j and let $f(z^j)$ be its density function. Consider $x^j = \alpha t \theta^j$ and $y^j = \tau(\alpha) \delta^j$. Clearly, x^j and y^j are also uniformly distributed random variables, with density $\frac{\psi}{\alpha t}$ and $\frac{\psi}{\tau(\alpha)}$, respectively. The extremes of the two variables are: $\underline{y}^j = \tau(\alpha)(-\frac{1}{2\psi} + \bar{\delta}^j)$; $\bar{y}^j = \tau(\alpha)(\frac{1}{2\psi} + \bar{\delta}^j)$, $\underline{x}^j = \alpha t(-\frac{1}{2\psi} + \bar{\theta}^j)$, $\bar{x}^j = \alpha t(\frac{1}{2\psi} + \bar{\theta}^j)$. In order to derive explicitly $f(z^j)$, assumptions are needed on the relative range of x^j and y^j . We assume through:

- A.1 $\tau(\bar{\alpha}) > \bar{\alpha}t$;

A.1 fits well the situation of our case study (and of many developing countries currently involved in a decentralization process), as transfers were by far the most important component of municipal financing in pre-reform Italy (see Section 3)⁹. Let $k = \frac{\psi^2}{\alpha t \tau(\alpha)}$. By the convolution theorem, under A.1, $f(z^j)$ is:

$$\begin{aligned} f(z^j) &= k(z^j - \underline{x}^j - \underline{y}^j), \text{ for } \underline{x}^j + \underline{y}^j \leq z^j \leq \bar{x}^j + \underline{y}^j; \\ f(z^j) &= k(\bar{x}^j - \underline{x}^j) = k \frac{\alpha t}{\psi} = \frac{\psi}{\tau(\alpha)}, \text{ for } \bar{x}^j + \underline{y}^j \leq z^j \leq \underline{x}^j + \bar{y}^j; \\ f(z^j) &= k(\bar{x}^j + \bar{y}^j - z^j), \text{ for } \underline{x}^j + \bar{y}^j \leq z^j \leq \bar{x}^j + \bar{y}^j; \end{aligned} \quad (3)$$

Secondly, we also impose an exogenous bound on the maximal difference between the two types' (expected) efficiency levels:

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

As can be easily checked, A.2 implies that $E(z^k)$ belongs to the "flat" part of $f(z^j)$, for $j, k = a, p$. This will be useful in simplifying the computations below.

2.1 The 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 $f(z)$ for both types. With the elections, period 1 ends and period 2 begins. Whoever in charge at the beginning of period 2, chooses again

⁹But notice that A.1 is not essential for the results to follow. As the forces we study would remain the same, it can be shown that the same results qualitatively occur even imposing the opposite assumption $\tau(\underline{\alpha}) < \underline{\alpha}t$. Details are available by the authors on request.

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 then realized. In both cases, g_2 is then determined. The game then 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). It has the advantage of greatly simplifying the analysis, ruling out signalling effects, while still providing electoral incentives to incumbent politicians. It may also have some bearing with reality, as it basically means that there is some common belief, shared by the incumbent j himself, about the ability of a fresh entrant of type j to administer municipal affairs, but that the true value of a politician will only be known after he has tried his hands in government. Notice that this assumption also implies that all politicians of type j , as they are all ex-ante identical, make the same choice of r_1^j in period 1. The assumption that, at the time he sets r_1^j , the incumbent j does not know the type of the opponent also seems very reasonable, as opponents are typically selected only few months before the elections. In any case, as will become clear as we proceed, relaxing this assumption would not affect much our results¹⁰. Notice that under the assumption that opponents are chosen randomly, this implies that at the time of setting r_1^j , incumbent j expects to face an opponent of type a with probability π and an opponent of type p with probability $1 - \pi$.

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

Having solved period 2, let us go back to period 1. At the end of this period, the voter observes g_1^j but she does not observe either r_1^j or the realization of z^j . The voter however expects the incumbent to take some rents in the first period (or the politician would deviate immediately and takes maximal \bar{R} in the first period too). We then look for an equilibrium where the voters use these expectations to discriminate between high/low quality incumbents. Let $r_1^{j^e}$ 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:

$$E(z^j | g_1^j) = \frac{g_1^j}{(1 - r_1^{j^e})} \quad (4)$$

Intuitively, a plausible strategy for the voter is then to vote for the incumbent if $E(z^j | g_1^j) \geq$

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

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

$E(z^k)$ and vote for the opponent k otherwise. We show below that this is indeed the optimal strategy for the voter. At the equilibrium, the incumbent knows the voter's optimal strategy when setting r_1^j , and knows $r_1^{j^e}$. Ex ante, he can then compute the probability of being reelected as a function of r_1^j , $r_1^{j^e}$, and of the expected type of the opponent. Using Eq. (2) and (4), and the timing assumptions above, the expected rents of an incumbent of type j over the two periods can be written as:

$$E(R^j) = r_1^j E(z^j) + \lambda \bar{R} \pi \left[1 - \text{prob} \left(z^j \leq E(z^a) \frac{(1 - r_1^{j^e})}{(1 - r_1^j)} \right) \right] + \quad (5)$$

$$+ \lambda \bar{R} (1 - \pi) \left[1 - \text{prob} \left(z^j \leq E(z^p) \frac{(1 - r_1^{j^e})}{(1 - r_1^j)} \right) \right]$$

Clearly, raising r_1^j increases expected rents in the first period, but for given $r_1^{j^e}$, it also reduces the probability of being re-elected in the second period, 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 to be confirmed, so we look for a solution of the incumbent's problem where $r_1^j = r_1^{j^e}$ also holds. Invoking Eq. (3) and A.2, deriving and imposing the equilibrium condition $r_1^j = r_1^{j^e}$, we get $r_1^{j^*}$, the equilibrium rents rate¹²:

$$r_1^{j^*} = r_1^{j^e} = 1 - \frac{\lambda \bar{R} \psi}{\tau} \left[\frac{E(z^o)}{E(z^j)} \right] \quad (6)$$

where $E(z^o) = \pi E(z^a) + (1 - \pi) E(z^p)$ ¹³. At $r_1^{j^*}$, expected rents in the first period for the j 's incumbent are then $r_1^{j^*} E(z^j) = E(z^j) - \frac{\lambda \bar{R} \psi}{\tau} E(z^o)$. Invoking Eq.(2), this implies that at the proposed equilibrium, the (expected) level of g_1 is just $g_1^* = \frac{\lambda \bar{R} \psi}{\tau} E(z^o)$ under both types of incumbent. It also follows that the more efficient incumbent, i.e., the incumbent with higher $E(z)$ ex ante, expects to get higher first period rents at the equilibrium. As shown, first period expected rents are instead decreasing in $\lambda \bar{R}$ (a larger $\lambda \bar{R}$ means that second period rents are either larger or that they matter more for the politician, and therefore he is willing to give up more current rents in order to be re-elected) and in the density $\frac{\psi}{\tau}$ (a larger $\frac{\psi}{\tau}$ means that the incumbent loses/gains more votes if r_1^j diverges from $r_1^{j^e}$). Note also that, at the equilibrium, a candidate j expects to be re-elected with probability $\frac{1}{2}$ if he meets a candidate of the same type, and to be re-elected with probability $\frac{1}{2} + \frac{\psi}{\tau} [E(z^j) - E(z^k)]$, $j, k = a, p$, $j \neq k$, if he meets a candidate of a different type¹⁴. Thus, at the proposed equilibrium, more ex ante efficient types also expect to be re-elected, and earn second period rents, with higher probability.

¹²See the Appendix for a formal proof.

¹³This assumes $r_1^{j^*} < 1$. Proposition 1 below provides conditions on \bar{R} that guarantees this to be the case.

¹⁴This follows from the fact that the relevant part of the distribution of z^j used to compute this probability is the "flat" part (see the Appendix). Note further that A.2 implies that the probability of being re-elected is strictly between 0 and 1.

The strategies of the voter at the proposed equilibrium are also straightforward. Substituting for r_1^{j*} in Eq.(4) and computing, 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, as this means that (under the expectation that j plays the optimal rent strategy r_1^{j*}) the realization of z for incumbent j has been at least as large as the expected realization of z for the opponent, $E(z^j)$. 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 very similar to the standard ones derived in this literature (Persson and Tabellini, 2000); the only 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.

Finally, for the proposed strategies to form an equilibrium, it must also be the case that even the weakest incumbent prefer to play this strategy rather than deviating and taking maximal rents in the first period (and not be re-elected in the second), and that the voter too prefers to play her proposed strategy rather than the alternative best strategy of always failing the weakest incumbent at the elections. As the Appendix proves, these conditions simply translate in upper and lower limits for \bar{R} . We can then conclude:

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

Proof. See Appendix. ■

2.2 Comparative statics

We now use our results above to study the effects of a decentralization reform on consumers' welfare and political candidacy in municipalities with a different tax base α . In the context of the model, this can be captured as an increase in t followed by a reduction in τ . In particular, in our case study, the Italian decentralization reform of the '90s, the introduction of the new property tax was accompanied by an offsetting variation in grants, so that at the statutory level of the new tax rate, each municipality had exactly the same resources both before and after the reform (see Section 3). In terms of our model, the simplest way to capture this

feature 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, $d\tau = -\alpha dt$. It also follows that $E(z^a) \geq (<)E(z^p)$ for $\alpha \geq (<)\alpha^*$ where $\alpha^* = \frac{\bar{\delta}^p - \bar{\delta}^a}{t(\bar{\theta}^a - \bar{\theta}^p + \bar{\delta}^p - \bar{\delta}^a)}$, provided that $\bar{\alpha} > \alpha^* > \underline{\alpha}$. 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.

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

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

Proof. See Appendix. ■

To provide an intuition for Proposition 2, note that a change in t , matched by a revenue 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 above on $\bar{\theta}^j$ and $\bar{\delta}^j$, this effect is certainly positive for the a -type, as $(\partial E(z^a)/\partial t - \alpha \partial E(z^a)/\partial \tau) dt > 0$, and certainly negative for the p -type, as $(\partial E(z^p)/\partial t - \alpha \partial E(z^p)/\partial \tau) dt < 0$. 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. This indirect effect is also negative (positive) in the second period for the p -type (a -type) as it reduces (increases) the probability of being re-elected. 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 in managing local resources, that is, on $\bar{\theta}^a$ being larger of some threshold, θ^* . But the important point, also stated in Proposition 2, is that regardless of its effects on the absolute level of politicians' utilities, the decentralization reform certainly makes the a -type better off relatively to the p -type, and particularly so in richer communities. This will be useful below.

What about the voter? In the first period, her 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 municipalities. 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 consumer living in a municipality with tax base

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

α , where $U^j(\alpha)$ is consumers' expected utility over the two periods when the first period incumbent is of type j , $j = a, p$. One can then show the following:

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

Proof. See Appendix. ■

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 municipality 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 are of a -type. This holds irrespective of the tax base, although in the latter case the consumers in rich municipalities gain the most from the reform. However, for intermediate values of π , it is the tax base that matters in determining the welfare effect of the reform. In particular, for given π , the richer is the municipality, 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 potential politicians. To study this case, suppose that at time 0, that is, before period 1 begins, a citizen of type j , $j = a, p$, is considering whether entering in the political arena or not. Suppose that there are n such potential candidates, where n is assumed to be a quite large number. The candidacy choice depends on the opportunity cost for entering in politics, that is on the remuneration that a potential candidate of type j could alternatively earn if she decided to remain a private citizen instead. Let us assume that the wage that each of the potential candidates j earns in the private market is drawn at the beginning of period 0 from a common independent uniform distribution on the interval $\{0, \bar{w}\}$. Citizen j observes the realization of her wage w^j before deciding whether becoming a member of the set of potential candidates j ; she also knows the expected two period rents for becoming an incumbent at time 1, $E(R^j)$. We assume that there are no costs in joining the set of potential candidates, and that both w^j and $E(R^j)$ are so large with respect to the benefits/costs that j receives from the municipality as a private citizen that she just ignores the latter in taking her candidacy decision. The only cost for a citizen j of becoming a politician is that if she is elected, she

has to rule, giving up her private wage. The candidacy decision is taken at the end of time 0 and cannot be revised afterward. After the candidacy choice has been taken and so the set of all possible politicians at the end of period 0 is determined, one candidate is chosen randomly by nature to become the incumbent politician in charge at period 1; the game then unfolds as already described in the previous Section.

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

Notice from the discussion above that while π depends on $E(R^j)$, $E(R^j)$ also depends on π , as the probability of being re-elected (and, therefore, second period rents of an incumbent) depends on the probability of meeting different types of opponents. Intuitively, higher expected rents for the more efficient type j induce more individuals of type j to enter the political market which in turn reduces expected rents, as it reduces the probability of meeting an opponent of a less efficient type. At the equilibrium π^* these two forces need to balance. As the Appendix shows, solving the resulting system of simultaneous equations and assuming an interior solution, this equilibrium share can be computed as:

$$\pi^* = \frac{1}{2} + \frac{(E(z^a) - E(z^p))(1 + \frac{\lambda \bar{R} \psi}{\tau})}{c(A + P)} \quad (7)$$

where $c = \frac{2\bar{w}(1+\lambda)}{n} > 0$. Thus, the denominator of the second term on the RHS of Eq. (7) is just proportional to the total number of potential politicians of both types and it is strictly positive. Eq. (7) allows us to get an important conclusion:

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

Hence, at the equilibrium with endogenous candidacy, richer municipalities have a larger share of a -type politicians and, therefore, in expected terms, a higher share of incumbent politicians of a -type¹⁶. Using Eq. (7), 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.

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

Proof. See Appendix. ■

Thus, the reform should have the effect of increasing even further the divergence between municipalities. After the reform, richer municipalities should have even more a -type politicians, while poorer municipalities 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. (7), as a -type incumbents become relatively more efficient than p -types. Under a very mild condition, discussed in the Appendix, the same reform would also have the effect of reducing the total number of politicians (the denominator in Eq. 7), as many p -type politicians would leave the political market and a -types may also leave the market (if $E(R^a)$ falls following the reform, see Proposition 2) or in any case the increase in their number is not enough to compensate for the exit of the p -types. Where $\pi^* \geq \frac{1}{2}$, the two effects work in the same direction, thus leading to an increase in π^* . Where $\pi^* < \frac{1}{2}$, but not too far from $\frac{1}{2}$, the first effect still dominates the second, so leading to an increase in π^* albeit at a reduced rate. Finally, for poor municipalities, the second effect may dominate, leading to a further reduction in π^* .

What would then be the effect of the reform on the welfare of voters? Endogenous candidacy clearly emphasizes what we already saw happening with exogenous politicians. Proposition 3 suggests that the tax reform 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 the reform should increase the share of a -type politicians more (or only) in the richer municipalities. Hence, in rich communities the reform should increase the expected welfare of voters for two reasons; directly, because it increases the utility of voters for given π , and indirectly, because it also increases π . On the contrary, in poor municipalities, 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 decentralization revenue compensated reform, like the Italian one, should have the effect of increasing the share of politicians with high administrative skills in rich municipalities. In poorer municipalities, the increase of the share of this type of politicians is smaller and might even become negative. Finally, the effect on voter's welfare also depends on the relative wealth of the municipalities. Consumers' welfare should increase in rich municipalities, while the effect on poor municipalities is uncertain and might even be negative. Bearing these predictions in mind, let us then examine our data.

3 A case-study: the Italian reforms

Municipalities play an important role in the Italian government system. They are in charge of a large number of services, ranging from general administrative services to local public good provisions (local transports, public parks and amenities, street lightening and cleaning, urban policy, sport policy, maintenance of school's building, kindergartens), from environmental services (garbage collection and disposal) to public utilities (heating and water

provision). Differences among the more than 8,000 municipalities are wide, with respect to size, average income, population density and composition.

At the beginning of the 90's, in the context of a more general decentralization policy that affected other levels of government and administrative units, two fundamental reforms were enacted on municipalities. First, in 1993, a new property tax (ICI), on the value of all buildings and lands was introduced (Legislative Decree 504/1992), providing Italian cities with a large and 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 allowances¹⁷. The central government compensated for the reform with a reduction in grants, so that at the minimum (compulsory) ICI tax rate, each municipality's revenues remained unchanged. As shown in Figure 1, the new tax had a dramatic effect on the composition of municipal revenues. For the Italian municipalities as a whole, the share of transfers in total revenues fell from just above 60% in 1992 to about 40% in 1993. However, the effect was differentiated across the country: in 2000, for instance, municipalities in the North were *on average* self-financed for about 70% of their budget, while in the poorer South grants covered *on average* about 60-70% of total municipal expenditure¹⁸.

The second reform concerned the electoral system. In the same year (Law 81/1993), the traditional parliamentary system was substituted by a quasi-presidential one, involving the direct election of mayors. Before 1993, citizens voted for parties' representatives to elect the city council, that then elected the mayor and the executive office. Since the reform, citizens directly elect the mayor, and a majoritarian prize guarantees that the parties supporting the winning candidate also get the majority of the seats in the city council¹⁹. 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 only²⁰. We exploit these institutional characteristics in our analysis below. Finally, notice that the mayor's salary (as well as the number of councilors and their salary) are determined uniformly across the country according to national rules²¹. They are related to the size of the municipality in terms of population, but not to local revenues (e.g., Gagliarducci and Nannicini, 2013).

It was not coincidental that both reforms were introduced simultaneously. They were

¹⁷Tax rates could be set in an interval between 0.4% and 0.7%, differentiating the rates between residential housing and commercial buildings. Municipalities could also introduce an allowance for resident house owners. See Bordignon et al. (2003) and Bordignon and Piazza (2010) for further details.

¹⁸See ISTAT, *I Bilanci consuntivi delle amministrazioni comunali*, June 2012, available at www.istat.it. Unfortunately, ISTAT does not provide a regional classification for all municipalities until 2000.

¹⁹There 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. (2010). Note, however, that all municipalities considered in our empirical analysis are above the 15,000 threshold, so that the electoral rules are the same.

²⁰The duration of the municipal legislature was 5 years before 1993, it was reduced to 4 years with the reform, and brought back to 5 years in 2000.

²¹In ordinary statute regions, which is why we only focus on these regions in the subsequent analysis.

part of the same decentralization strategy, that attempted to make local governments both more financially responsible, by providing them with larger revenue autonomy, and more accountable, by letting the executives be directly elected by citizens²². It is also clear that after the reform, as pointed out by a large literature in political sciences (e.g., Pasquino, 2006; Baldini and Legnante, 2000; Bettin Lattes and Magnier, 1995), Italian mayors assumed a paramount role in municipal policy. This justifies our focus on mayor's characteristics in what follows.

4 The empirical analysis

4.1 Identification strategy

In 1993 Italian municipalities were thus (simultaneously) subjected to two reforms. One, the reform of the electoral rules, affected all municipalities in the same way. We should then expect this reform to have had the same effect across municipalities on the local political class. The other, the reform of the funding system, affected the municipalities differently, depending on their tax base. On the basis of our theoretical model, we should then expect this second reform to have had different effects both on the selection of politicians and on citizens' welfare.

To test these differential impacts, we consider a simple difference-in-differences strategy, similar to the approach proposed by Card (1992)²³. In particular, we estimate the 'average treatment effect' (*ATE*) of the funding reform of municipalities, where the 'treatment' is represented by the reduction in the degree of VFI following the introduction of ICI. Let's define with Y^j ($j = a, p, w$), the variables indicating the a -type politicians, p -type politicians and consumers' welfare, respectively. As data on the ICI tax bases are likely to be endogenous²⁴, we use instead the per-capita Value Added at the provincial level averaged over the pre-treatment years 1991-1992, as an indicator of the level of autonomous resources of a municipality²⁵. On the basis of our theoretical model, we expect *ATE* to be positive for both Y^a and Y^w , and negative for Y^p . Identification of these effects is based on the large variation in income distribution across Italian municipalities²⁶. Our empirical model is represented by the following equation:

$$Y_{it}^j = \alpha \mathbf{A}_i + \beta \mathbf{B}_t + \gamma \mathbf{X}_{it} + \delta DECENTR_t \times VALUE_ADDEDpc_i + \varepsilon_{it} \quad (8)$$

²²Similar reforms were also introduced in the same period for the other two sub-levels of governments, regions and provinces.

²³See also Angrist and Pischke (2009) chapter 5.

²⁴Municipalities are in charge of building permits and in determining the use of buildings and lands, and this clearly affects the ICI tax base.

²⁵Data are provided by the "Istituto Tagliacarne" (<http://www.tagliacarne.it/>). In the WP version of this paper, we considered provincial level per capita GDP averaged over the years 1995-2010 from the same source. We decided to use the average of per capita Value Added over the pre-treatment years 1991-1992 in order to avoid endogeneity problems potentially affecting also income per capita. Notice that the correlation between our measures of average per capita Value Added and per capita GDP is 0.9487. Unsurprisingly, results discussed below are basically unaffected by this choice.

²⁶In our sample, the range of per-capita Value Added goes from 8,851 to 26,041 euro.

where \mathbf{A} and \mathbf{B} are fixed effects for municipalities and years, \mathbf{X} is a vector of political and economic variables to characterize differences across municipalities, $VALUE_ADDEDpc$ is the municipal per capita Value Added and $DECENTR$ is a dummy variable to identify the year of the ‘treatment’²⁷. Our hypotheses are then tested by looking at the sign and the statistical significance of δ , which capture the differential responses across municipalities generated by the national reform according to the different levels of wealth and, thus, fiscal autonomy.

4.2 Data and variables definition

4.2.1 Dependent variables

Type. The first prediction from our theoretical model suggests that a decentralization reform should change the share of politicians with administrative/political skills differently in rich and poor municipalities. To define politicians type we build an original database which contains detailed information on the mayors of the 89 chief provincial towns of the 15 ordinary statute regions elected between 1988 and 1997, the ten years period around the introduction of ICI. The database has been built starting from the archive provided by the Italian Ministry of Internal Affairs. The archive includes information on sex, age, date and place of birth, party affiliation, level of education, and the profession of the mayor before entering in politics. We exploit these data to define our proxies for the skills of local politicians. Notice that despite its usefulness, information in this archive are based on self-declaration by the same politicians. Hence, original information have been carefully revised using other sources, such as newspaper’s online archives, internet sites, and telephone calls to municipalities. For instance, we identify the exact party affiliation of each mayor, as this information is often replaced in the archive by other labels (e.g., civic lists) that did not allow us to identify precisely the mayors’ political coalition.

The most difficult part of our exercise is to find – starting from available information – reasonable proxies for the different skills of politicians in order to identify *a*- and *p*-types. The notion of “administrative” skills is obviously ‘fuzzy’, particularly when referred to local government officials (e.g., Perry,1989). To overcome this difficulty, we use the *profession of the mayor before entering in politics* as a proxy for his administrative skills, and we classify these professions according both to the job classification provided by the Italian National Institute of Statistics (ISTAT), and to an index of skills required for each profession provided by the Italian Institute for the Training of Workers²⁸. In particular,

²⁷We use the first year an election took place for each municipality with the new electoral rules to identify the year of the treatment. In fact, although some municipal elections took place already in 1993, for most municipalities the new rules applied with some delay, as the previous legislature finished later. Notice that this also allows us to identify the impact of decentralization on Y^w when defined using the amount of separate waste collected, data for which are available only starting from 1993.

²⁸The index provided by ISFOL is built starting from a survey in which an individual is asked to rate (from 0 to 100) how important is a certain skill for her job, and how complex is a certain task for which this skill is required. We consider in particular the two skills "manage" and "decide" to define the "administrative" skills. Notice that for certain jobs, like for biologists, the index appears to be disproportionately high, since it is affected by the presence of directors in hospitals. We combine the ISFOL index with the ISTAT

in the following, we propose and test the model using three different definitions of Y^a . Definition [A] is the broadest and coincides with the set of “top” private professions. It includes the following jobs: entrepreneurs, managers and directors of private and public firms, engineers and architects, business consultants, lawyers, university professors²⁹. We exclude business consultants and lawyers from definition [B], on the basis of the argument that their skills are perhaps more useful for the private sector than the public one. Finally, we further exclude university professors³⁰ from definition [C], that therefore only includes entrepreneurs, managers, engineers and architects. Y^a is then a dummy that assumes value 1 when the mayor belongs to one of the categories included in definitions [A], [B] and [C], depending on the classification used, and 0 otherwise.

As for the p -type politicians, we proxy “political” skills with the *previous political experience of the candidate* before becoming mayor. We use several sources to identify all the political offices previously held by a mayor³¹, including all legislative and executive positions in all local governments (regions, provinces, municipalities) and the Italian and European Parliament. We then define the variable Y^p as the ratio between the years of previous political experience and the working age of the mayor (i.e., his age minus 17). Hence, Y^p is a continuous variable measuring the percentage of working years a candidate had already dedicated to politics before becoming mayor.

Performance. The second prediction from our theoretical model suggests that a decentralization reform should impact on consumers’ welfare differently in rich and poor municipalities. Like skills, welfare is also difficult to measure, so that we proxy here Y^w using two variables informative of the ex-post performance of mayors, that in our model is akin to voters’ welfare. According to definition [A], Y^w is the percentage of *separate waste collection*, an indicator provided by Legambiente, an Italian independent environmental organization. Managing waste is an important task assigned to municipalities, on which mayors are easily evaluated by citizens and for which they are considered accountable³²; it is often used by international organizations (e.g., the United Nations Habitat Programme) as a measure of the quality of municipal governance. Notice that this indicator has the further advantage of being available for each municipality and each year since 1993, and to be more reliable as an indicator of good governance for Italian municipalities than budget data³³.

According to definition [B], Y^w is the *probability of completing the term in office* (i.e., of *not* having an early termination, which we measure following Gagliarducci and Paserman,

classification, and just consider directors among our definition of Y^a . As a separate exercise, however, we also used the ISFOL classification as a proxy for administrative skills in our empirical analysis below, getting remarkably similar results. These are available from the authors on request.

²⁹It should be noted that not all these professions require a formal education. For instance, in 90’s Italy, most entrepreneurs and managers did not hold a college degree.

³⁰Since the category includes a wide range of disciplines, it is unclear whether university professors have the skills that could be useful to municipalities.

³¹We consider in particular: the online registry office of the Italian Ministry of Internal Affairs for all local levels of government; the online historical archive of both the Italian Chamber of Deputies and the Senate of the Italian Republic; the online archive of the European Parliament.

³²See, for instance, the echo in international media caused by the most recent rubbish crisis in Naples in 2008, and in Palermo in 2012.

³³Recall that identification of the impact of decentralization is allowed here by the misalignment in electoral years across municipalities. See footnote 25 above.

2012). Data are taken from the administrative archive of the Ministry of the Internal Affairs. In particular, we consider as cases of failed legislature only those following the resignation of the mayor, the resignation of the majority of the council, or a no-confidence vote in the council. Other technical reasons for early termination, like the death of the mayor, are not informative of the performance and are coded as completion of the term. We take this variable to be a direct evaluation by citizens of the performance of the mayor³⁴.

Descriptive evidence. Table 1 reports the unconditional means of all our dependent variables before and after the introduction of ICI, and the differences in means by Value Added quartiles. In general, this preliminary evidence provides support to our hypotheses. In particular, starting from Y^a , we find that the increase in the share of a -type politicians was larger in richer municipalities than in poorer ones according to two out of three definitions. In fact, according to definition [A], politicians with high administrative skills increased by 22.93% in the richest communities, while just by 8.06% in municipalities belonging to the bottom quartile. Definition [B] of Y^a roughly provides the same story (with an increase by 19.68% and 1.41% for the top and the bottom quartiles, respectively), while the evidence is much less clear with definition [C] of Y^a . As for Y^p , while the percentage of political experience of mayors increased for municipalities in the bottom quartile by 2.77, we register a sharp decrease for communities in the second and the third quartiles (by about 9.03 and 12.37 respectively), and still a decrease for the richest municipalities (by 3.72 percentage points). Finally, we find that the increase in one of our measures of consumers' welfare is coherent with the prediction of our theoretical model. According to definition [A] of Y^w , we observe an increase in the amount of separate waste collected in all municipalities, but this ranges from 0.44% in those belonging to the bottom quartile to 3.12% in the richest ones. As for the probability of completing the term in office (Y^w , definition [B]), on the contrary, the evidence points to a larger increase for municipalities in the bottom quartile with respect to those in the top one.

[Table 1 about here]

4.2.2 Control variables

Vector \mathbf{X} include a number of covariates that could also affect the selection of politicians. Descriptive statistics for all these variables are in the Appendix Table A.1.

Individual characteristics. We control for both the age and the gender of the elected mayor, considering the variables AGE and the dummy variable SF , which takes value one when the mayor is a female.³⁵

³⁴Resignation could also be a strategic option for mayors, say, in order to run for re-election and obtain a greater majority. However, as discussed in Gagliarducci and Paserman (2012), this does not seem to be the case in Italy. First, the two-term limit, the cost of campaigning, and the possibility to lose the mayoral monthly salary act to reduce the incentives for mayors to resign voluntarily; second, in the case of early termination (and no binding term limit), the available evidence suggests that not all mayors run for re-election and those who succeed in being re-elected are not able to significantly increase the size of their majority; finally, government instability has a clear negative connotation.

³⁵These two variables could be 'bad controls' in the terminology of Angrist and Pischke (2009), i.e.

Political controls. We include in the model a number of variables catching political characteristics of the mayor and of the municipality. As for the former, we first include two dummy variables, *LN* and *FI*, which take value one when the mayor belongs to *Lega Nord* (Northern League) and to *Forza Italia* (Forward Italy), respectively. These are the two new parties that substituted, at the beginning of the '90s, the old Italian parties on the center right of the political spectrum (Socialists and Christian Democrats) after a number of scandals led to their disappearance. It is then important to control whether the effect on the selection of politicians was due to these new parties, that could potentially rely on a different pool of candidates than the old ones. We also include in the model considering citizens' welfare, two additional variables (and their interaction) with the aim of separating 'selection' from 'incentives'. The first one is a dummy taking value one if the mayor is at her/his second term in office (and therefore cannot run again, i.e., is term limited, *TL*). Albeit rough, this is a control allowing us to separate the 'selection' effect from the 'disciplining' effect of elections, and it has been often used in the literature to this aim, including in the Italian case³⁶. The second one is a variable measuring the number of years in office since the last election (*CYCLE*). The idea here is to capture the political budget cycle, which can also affect the 'disciplining' effect of elections (e.g., Rogoff and Siebert, 1988).

As for the political characteristics of the municipality, we first consider the dummy variable *ALIGNMENT*, which takes value one when the mayor's party is the same as the one in power at the central level (in order to capture the impact of political alignment on mayors' election). Since competition may have an important impact on the selection of politicians (e.g., Galasso and Nannicini, 2011, for the Italian case), we finally consider an index of political competition, defining the variable *COMPETITION* as the Herfindahl index on the shares of votes taken by all the Italian parties active at the municipal level³⁷. To avoid endogeneity, we build this index by using data from the European elections at the municipal level for the closest round of European election held before the municipal election.

Socioeconomic controls. We include in the model the variable *POP* measuring the number of citizens in each municipality. This allows us to control both for the fact that a larger number of citizens allows for a larger pool of potential candidates, and for the incentive provided by the mayors' wage, which increases according to population size (e.g., Gagliarducci and Nannicini, 2013). To take into account the characteristics of the potential pool of candidates that can impact on the type of jobs and the political career, we control for the share of people older than 65 years out of the total municipal population (*OVER65_PERC*). We also introduce *MOVIPROV*, the number of enterprises relative to total population at

outcome variables of our 'experiment'. Consequently, we tested the robustness of our findings to their inclusion in the model. As they do not have any impact on our results, we decided - coherently with the literature - to retain both in our specifications.

³⁶The idea is that a mayor in his last term should not care (or should care less) about the 'disciplining effect' as he cannot be elected again. Hence, if a relation still exists between the quality of a politician and the quality of policy, this should be due to the 'selection effect' only. See, e.g., Besley and Case (1995) for the general argument and Bordignon et al. (2003) for an application to the Italian case.

³⁷In the last 20 years the Italian political system has been basically bi-partisan, with two large coalitions of parties, one on the Center-Right and one on the Center-Left, competing at national and local elections. See Bordignon et al. (2010) for further details.

the provincial level, as a measure of the opportunity cost of entering into politics.

5 Results

5.1 Baseline results

We estimate Equation (8) with OLS, using two-way cluster-robust standard errors to account for both serial and panel correlations (Cameron et al., 2011). Estimates of our coefficient of interest δ are reported in Table 2 for all the dependent variables Y^a , Y^p and Y^w .³⁸ For each of these, the baseline model in col. I (which includes also the variable *CYCLE* in the consumers' welfare specification) is augmented considering the two political variables *ALIGNMENT* and *COMPETITION* in col. II, and further augmented with additional political controls (the two dummies for *Lega Nord* and *Forza Italia*, and the dummy for term limited mayors in the model for consumers' welfare) in col. III. The estimated $\hat{\delta}$ take up the expected signs (positive for Y^a and Y^w , negative for Y^p), and are statistically significant at the usual confidence levels for all the six outcome variables, but for Y^a defined according to the more stringent definition [*C*]. (In this last case, the coefficient is never statistically significant, but it retains the positive expected sign). Notice also that the magnitude of the estimated coefficients is stable across specifications: as for the share of *a*-type politicians, *ATE* ranges from 0.135 to 0.156 in the case of Y^a definition [*A*], and from 0.110 to 0.140 in the case of Y^a definition [*B*]; as for *p*-type politicians, *ATE* is between -0.070 and -0.086; finally, in the case of consumers' welfare, the effect is from 1.608 to 1.961 using separate waste collection, and from 0.171 to 0.189 in the case of the probability of completing the term in office.³⁹

[Table 2 about here]

As is well known, difference-in-differences estimates relies on the assumption of a common trend for all municipalities. In order to test whether our results depend on the funding reform, and are not biased by differences in pre-existing time trends, we then run a set of placebo tests using only data from the pre-treatment years and introducing a fake funding reform. In particular, for the dependent variables Y^a , Y^p and Y^w definition [*B*] we consider two experiments over the period 1988-1992, estimating Equation (8) using two different placebo treatments: 1) $F_DECENTR_{t1} \times VALUE_ADDEDpc_i$ equal to 0 before 1990 and to the per capita Value Added from 1990; 2) $F_DECENTR_{t2} \times VALUE_ADDEDpc_i$ equal to the per capita Value Added starting from the first municipal election held after 1989 (and 0 otherwise). As for Y^w definition [*A*] - since data before 1993 are unavailable - we run

³⁸The complete tables for all specifications in the paper are available in the latest working paper version.

³⁹As for the other controls, a persistent and interesting result is the role of new political parties in influencing the choice of candidates. Both *Lega Nord* and *Forza Italia* picked up more *a*-type and less politically experienced candidates. Also interesting is the variable *CYCLE* in the separate waste collection models: the positive coefficient suggests the presence of a political budget cycle. Finally, the term limit does not have any impact on our indicators of welfare, pointing to the importance of a 'selection effect'. See the complete tables for details.

a placebo experiment on the years 1993-1994 only for the municipalities that voted for the first time after the introduction of ICI starting from 1995. In this case the fake treatment $F_DECENTR_{t3} \times VALUE_ADDED_{pc_i}$ becomes equal to the per capita Value Added in 1994. The results of these placebo experiments are reported in Table 3. None of the coefficients is statistically significant, and some of them even present the opposite sign with respect to the one expected. This reassures us that our results are not biased by differential time trends among different groups of municipalities.

[Table 3 about here]

5.2 Robustness checks

Our baseline results are subject to a number of additional concerns, that need to be accounted for. These include the interaction of the decentralization process with other political variables, the role of the electoral reform, the importance of political campaign costs, the role of social capital, the definition of Y^a and Y^p in accordance to our theory. In what follows we explore each of this in turn. The general conclusion is that – despite the strong requirements imposed to the model for parameters identification – our findings appear to be remarkably robust, and strongly support our story about the selection of different politicians in the presence of different degrees of VFI.

The interaction of political variables with decentralization. A first concern is that the decentralization process may also have interacted with other political variables, so that our explanation may miss some important elements of what happened in municipal policy after the reform. To check for this, we augment the models in Table 2 col. III with additional interaction terms for all the political variables (*ALIGNMENT*; *COMPETITION*; *LN*; *FI*; and *CYCLE* and *TL* in the Y^w equation), considering interactions with both *VALUE_ADDED_{pc}* (col. IV) and *DECENTR* (col. V). Our previous findings remain valid in the case of interactions with *VALUE_ADDED_{pc}*, in terms of both magnitude and statistical significance. On the contrary, most coefficients turn insignificant in the case of interactions with *DECENTR*. However, this is mainly due to the almost perfect collinearity between our variable of interest and the interaction between *COMPETITION* and *DECENTR*, as is testified also by the fact that the standard error for $\hat{\delta}$ (and the coefficient itself) blows up in this specification. In the only two cases in which the coefficient remains statistically significant (in the Y^p and the Y^w definition [A] equations), the magnitude almost double with respect to specifications in col. I-IV.

The role of the electoral reform. A second and perhaps more important problem with our results is that the tax decentralization was implemented in the same year in which the electoral reform was implemented. There remains therefore the suspicion that our results on the selection of politicians are really driven by the electoral reform, that for some reasons affected differently rich and poor cities, rather than by the change in the financing system. To address this issue, we then consider a second tax decentralization reform that occurred in the '90s, this time at *unchanged electoral rules*. In 1999, municipalities were also given

the opportunity of introducing a surcharge (up to 0.5%) on the tax base of the personal income tax (PIT) of their residents; a possibility that they had up to 2002, when the central government decided to freeze the surcharge⁴⁰. Differently from the municipal property tax, the introduction of the PIT surcharge was not matched by a contemporaneous reduction in central government transfers (because the surcharge was not mandatory), and it was in any case less powerful in terms of the potential reduction of VFI⁴¹. Still, it is worth investigating if even this smaller reform had effects on the selection of politicians and consumers' welfare.

To this aim, we estimate our previous Equation (8), where the dummy *DECENTR99* takes now value one from 1999 onwards, on an extended dataset relative to the period between 1995 and 2002. Table 4 replicates the same models in previous Table 2. In this case, the coefficient $\hat{\delta}$ turns out to be significant in all the five specifications and with the expected sign only in the Y^a equation, when considering the more stringent definition [C]. It is insignificant considering larger definitions [A] and [B] of the a -type politicians, but the sign is the expected one. For the Y^p equation, $\hat{\delta}$ presents the expected sign and it is significant in the first three specifications; it becomes insignificant - though with the expected sign - in the last two models. As for consumers' welfare, $\hat{\delta}$ is significant in the first four specifications when considering definition [A] of Y^w , while it is insignificant when using definition [B].⁴²

[Table 4 about here]

The importance of political campaign costs. A third source of concern is related to political campaign costs (e.g., Poutvaara and Takalo, 2007). As an alternative explanation, one could argue that following the reform the stronger emphasis on the role of mayors may have led candidates to spend more money on their campaigns, and as richer people can spend more, this could explain why we observe more mayors coming from top positions in the private sector in the richer cities⁴³. We cannot directly control for this issue, as data on campaigning costs at the municipal level are not publicly available. But one could argue that this phenomenon should have been particularly relevant in rich metropolitan cities, such as Milan or Rome, rather than say, in Cremona and Viterbo, two small provincial towns belonging to the same regions. We then re-run all our regressions of the model in column III of tables 2 and 4, dropping the 15 cities that are also regional capitals (*Capoluogo di Regione*), typically the richest ones. As shown in Table 5, our previous results are all confirmed. In the first period, characterized by the introduction of ICI, *ATE* is still significant and with the expected sign in all the specifications but in the Y^a equation when

⁴⁰See Bordignon and Piazza (2010) for details. The municipal PIT surcharge was reintroduced in 2007.

⁴¹In the sample considered by Bordignon and Piazza (2010), for instance, the PIT revenues in 2002 was about 8% of total tax revenues of municipalities, in contrast with about 40% for ICI revenues.

⁴²As for the other variables, it is interesting to note that the role of new political parties in selecting new types of politicians was apparently limited to the period following the introduction of ICI. On the contrary, we still find evidence of a political budget cycle. Finally, the term limit is again insignificant in all specifications. See the complete tables for details.

⁴³To be sure, this is less of a concern in Italy than in other countries, as political parties usually finance the electoral campaigns of their candidates, and political parties are in turn generously financed with public funds.

using definition [C]. In the second period, with the introduction of the PIT surcharge, the estimated impact of the reform is significant and with the expected sign only when using definition [C] in the Y^a equation, in the Y^p model and in the Y^w equation when considering separate waste collection.

[Table 5 about here]

The role of social capital. Social (or civic) capital is generally thought of as being very different between the Centre-North and the South of the country, and to have an impact on a wide range of economic and political variables (e.g., Guiso et al., 2011). Hence, it might also have played a role in the selection of politicians and on consumers' welfare following the decentralization reform. As social capital is a long run phenomenon, the inclusion of municipality fixed effects \mathbf{A} in our specifications should already work as a control for the different level of social capital across Italian cities⁴⁴. However, as an additional check, we decided to consider separately the two sub-samples of Centre-Northern and Southern provinces. If our theory is correct, income variability *within each area*, characterized by similar levels of social capital, should induce the same process of selection we detected in the whole sample. Results taking this route, considering previous model in col. III of Tables 2 and 4, are presented in Table 6. As shown, in the Center-North, where the income differential across municipalities is still large⁴⁵, we find supporting evidence for our story, especially in the case of ICI. All coefficients have the expected signs and are in most cases statistically significant. Results for the Southern municipalities are instead less clear-cut. In particular, we find the expected negative and significant coefficient for Y^p , and the expected positive and significant coefficient in the equation modelling the probability to complete the term in office, but only in the case of the ICI reform. The impact is insignificant in all the remaining equations. Possibly, these results suggest that richer municipalities in the South also witnessed a change in the local political class in the aftermath of the reforms, but that they were less successful in attracting politicians from the private sector than their Northern counterparts⁴⁶.

[Table 6 about here]

The definition of Y^a and Y^p . A final issue concerns the empirical definition of our indicators Y^a and Y^p . In our theoretical model the a -types and p -types are two mutually exclusive types of politicians, while in the empirical analysis so far we “mixed up” the

⁴⁴Re-running our regressions introducing the traditional indicators of social capital used in the empirical literature (like blood donations or participation to political referenda) would not help, as these indicators are typically available for few years only and do not present large variability across periods, i.e. they are akin to fixed effects.

⁴⁵In our sample, the average per-capita Value Added was 18,103 euro for municipalities in the Centre-North, varying in the interval between 12,531 to 26,040 euro. In the South the average was 11,924 euro per capita, varying in the interval between 8,851 to 16,301 euro.

⁴⁶These are of course just potential explanations given the limitations of our data set, as in the South the number of municipalities included in the sample is small and the range of income differential is also much smaller.

two types, focusing rather on the skills of the different politicians. Thus, for example, a mayor coming from a top private profession would be considered an a -type even if, say, she/he also had a long experience in politics. To check if this affects our results, we need to provide a definition of Y^a and Y^p which is mutually exclusive. One way to proceed is to restrict the definition of p -types further, re-defining Y^p as a dummy variable which takes value 1 when the mayor has devoted more than half of her working years to politics. This allows us to define: (pure) a -types politicians as those for which $Y^a = 1$ and $Y^p = 0$; (pure) p -types politicians as those for which $Y^a = 0$ and $Y^p = 1$; ‘impure’ types, i.e., those cases for which a precise assignment of the mayor to one of the types could not be made, have been coded as zeros, *de facto* introducing stronger requirements for parameters identification. We then re-estimate previous model in Tables 2 and 4 col. III considering this more stringent definitions of the two types. Results in Table 7 again broadly confirm our story for both instances of financing reforms, ICI and PIT surcharge, at least for a -types: In the richest cities, decentralization caused a statistically significant increase in the share of mayors with administrative skills with respect to poorest municipalities. Coefficients are instead statistically insignificant for p -types, even though signs are mostly negative as expected.

[Table 7 about here]

6 Concluding remarks

This work supports the idea that decentralization requires, in order to work properly, enough autonomous resources at the local level. But while the literature usually points to a “common pool” effect as an explanation for the observed negative relationships between VFI and decentralization outcomes, we suggest here that there may even be a “selection effect” on local politicians, finding strong supportive evidence for this argument in the case of the Italian decentralization reforms of the ’90s. Intuitively, there is no reason for voters to choose a good administrator of local matters when the basic task of the local politician is to ‘bribe’ the center in order to guarantee that central money keeps flowing to the community; political skills dominate. In principle, of course, 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 type of transfers (for instance, capital expenditure grants). As some political discretion is therefore unavoidable, in those countries and situations where VFI is large, it is then likely that decentralization may meet the kind of problems we discuss in this paper.

Finally, our findings can also be read as a criticism to the economic literature that, building on the seminal work by Putnam (1993), focuses on “social capital” as an explanation for the different success of decentralization experiments. On these grounds, note that our

results hold even when we control for proxies of “social capital” indicators and even when we split the sample to consider municipalities belonging to areas with different endowments of social capital. This suggests that while history and culture certainly matter, one should not underestimate the role of rational behavior on the part of voters in determining outcomes.

ACKNOWLEDGEMENTS. *We wish to thank Alessandro Fontana, Vincenzo Galasso, Anke Kessler, Marko Koethenburger, Giovanna Messina, Tommaso Nannicini, Giacomo Ponzetto, Francesco Porcelli, Panu Poutvaara, Marzia Romanelli, Paolo Sestito, Francesco Sobbrío, Guido Tabellini, Davide Ticchi, Dave Treisman, Andrea Vindigni, Jingjing Zhang, and all seminar participants at the Bank of Italy Local Governments Workshop, the CESifo Public Sector Economics Workshop, the Mills Workshop, the 68th IIPF Congress (Dresden University of Technology), the EPCS 2013 Meeting (ETH-KOF, Zurich), the 5th Workshop on Fiscal Federalism (Universitat de Barcelona), University of Milan Bicocca, University of Trieste, University of Tor Vergata and the Institute for Advanced Studies (Lucca), for useful comments. Usual disclaimers apply.*

References

- [1] Angrist J.D. and Pischke J-S (2009). *Mostly Harmless Econometrics*. Princeton University Press.
- [2] Ahmad E., Brosio G., eds. (2008). *Handbook of Fiscal Federalism*. Edward Elgar, London.
- [3] Baldini G., Legnante G. (2000). *Città al voto: i sindaci e elezioni comunali*. Il Mulino, Bologna.
- [4] Bardhan P., D. Mookherjee (2000). “Capture and Governance at Local and National Levels.” *American Economic Review*, 90 (2): 135-139.
- [5] Bardhan P., D. Mookherjee, eds. (2006) *Decentralization to Local Governments in Developing Countries: A Comparative Perspective*, MIT Press.
- [6] Besley T. (2004). “Paying politicians: theory and evidence.” *Journal of the European Economic Association* 2, 193-215.
- [7] Besley T. (2005). “Political Selection.” *Journal of Economic Perspectives*, 19(3): 43–60.
- [8] Besley T. (2006). “Principled Agents? The Political Economy of Good Government”. Oxford University Press, Oxford.
- [9] Besley T., A. Case (1995). “Incumbent Behavior: Vote Seeking, Tax Setting and Yardstick Competition”, *American Economic Review*, 85(1): 25-45.

- [10] Besley T., Smart M. (2007). "Fiscal restraints and voter welfare", *Journal of Public Economics*, 91(3-4): 755-773.
- [11] Bettin Lattes G., Magnier A. (1995). "I nuovi sindaci: come cambia una carriera politica". *Rivista Italiana di Scienza Politica*, 25(1):91-118.
- [12] Bordignon M., Piazza S. (2010). "Who do you Blame in Local Finance? An Analysis of Municipal Financing in Italy". CESifo Working Paper, n. 3100.
- [13] 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.
- [14] Bordignon M., Colombo L., Galmarini U. (2008). "Fiscal federalism and lobbying". *Journal of Public Economics*, 92(12): 2288-2301.
- [15] Bordignon M., Nannicini T., Tabellini G. (2010). "Moderating political extremism", mimeo, Bocconi University.
- [16] Brollo F., Nannicini T., Perotti R., Tabellini G. (2013). "The Political Resource Curse". *American Economic Review*, 103 (2013), 1759-1796.
- [17] Cai H. and Treisman D. (2005). "Does Competition for Capital Discipline Governments? Decentralization, Globalization and Public Policy." *American Economic Review*, 95(3): 817-830.
- [18] 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.
- [19] 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.
- [20] Caselli F. and Morelli M. (2004). "Bad politicians". *Journal of Public Economics*, 88, 759-782.
- [21] 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.
- [22] Eyraud L., Lusinyan L. (2011). "Decentralizing Spending More than Revenue: Does It Hurt Fiscal Performance?", IMF Working Paper, n. 226.
- [23] Ferraz C., Finan F. (2009). "Motivating Politicians: the impacts of monetary incentives on quality and performance." NBER Working Paper No. 14906.
- [24] Fisman, R. and R. Gatti (2002), "Decentralization and Corruption: Evidence from US Federal Transfer Programs", *Public Choice*, 113: 25-35.

- [25] Gagliarducci S., Nannicini T. (2013). "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection". *Journal of the European Economic Association*, 11: 369-398.
- [26] Gagliarducci S., Nannicini T. and P. Naticchioni (2010). "Moonlighting Politicians" *Journal of Public Economics*, 94(9-10), 688-699.
- [27] Gagliarducci S., Paserman D. (2012). "Gender Interactions within Hierarchies: Evidence from the Political Arena" *Review of Economic Studies*, 79(3), 1021-1052.
- [28] Galasso V., Nannicini T. (2011). "Competing on good politicians", *American Political Science Review*, 105(1): 79-99.
- [29] 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.
- [30] 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.
- [31] Hines J. R. and R. H. Thaler (1995). "The Flypaper Effect". *Journal of Economic Perspectives*, 9(4): 217-226.
- [32] Lockwood B. (2008). "Fiscal Decentralization: A Political Economy Perspective", in E. Ahmad and G. Brosio (eds.), *The Handbook of Fiscal Federalism*, Edward Elgar (revised version 2013, in press).
- [33] Magnier A. (2004). "Between Institutional Learning and Re-legitimization: Italian Mayors in the Unending Reform". *International Journal of Urban and Regional Research*, 28(1): 166-182.
- [34] Mattozzi A. and Merlo A. (2008). "Political careers or career politicians?" *Journal of Public Economics*, 92(3-4): 597-608.
- [35] Pasquino G. (2006). *I sistemi elettorali*. Il Mulino, Bologna.
- [36] Perry J. L. (1989). "The effective public administrator", in J.L. Perry and Associates (eds.), *Handbook of Public Administration*. San Francisco, Jossey-Bass Publishers.
- [37] Persson T., Tabellini G. (1994). "Does centralization increase the size of government?". *European Economic Review*, 38(3-4): 765-773.
- [38] Persson T., Tabellini G. (2000). "Political Economics: Explaining Economic Policy". Cambridge, MA: MIT Press.
- [39] Poutvaara P., Takalo T. (2007). "Candidate quality". *International Tax and Public Finance*, 14: 7-27.

- [40] Putman R. (1993) *Making Democracy Work*. Civic tradition in Modern Italy. Princeton University Press.
- [41] 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.
- [42] Rodden J. (2006). "The Political Economy of Federalism" in Weingast B. and D. Wittman (eds.), *Oxford Handbook of Political Economy*, Oxford University Press.
- [43] Rodden, J. A., Eskeland G. S., and Jennie L. (2003). "Fiscal Decentralization and the Challenge of Hard Budget Constraints." Cambridge, MA: MIT Press.
- [44] Rogoff K., Siebert A. (1988). "Elections and Macroeconomic Policy Cycles". *Review of Economic Studies*, 55:1-16.
- [45] Treisman D. (2007). "The Architecture of Government. Rethinking Political Decentralization", Cambridge: Cambridge University Press.
- [46] 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.

Tables and figures

Table 1. Mean values of dependent variables by Value Added pc quartiles (1988-1997)

| Quartile ^c | Type ^a | | | Yp | Performance ^b | |
|-----------------------------------|-------------------|-------|-------|--------|--------------------------|-------|
| | Ya | | Yw | | [A] | [B] |
| | [A] | [B] | | [C] | | |
| All years | | | | | | |
| First | 44.24 | 19.35 | 11.98 | 19.65 | 0.89 | 73.27 |
| Second | 26.92 | 15.87 | 12.50 | 25.00 | 6.49 | 83.33 |
| Third | 36.82 | 20.45 | 10.45 | 23.79 | 8.87 | 84.47 |
| Fourth | 35.81 | 19.53 | 11.16 | 27.10 | 8.56 | 80.37 |
| Before ICI | | | | | | |
| First | 41.30 | 18.84 | 10.14 | 18.65 | 0.57 | 60.74 |
| Second | 22.31 | 13.22 | 13.22 | 28.60 | 4.68 | 73.64 |
| Third | 23.88 | 11.94 | 11.94 | 28.63 | 8.49 | 77.44 |
| Fourth | 27.07 | 12.03 | 11.28 | 28.49 | 6.14 | 69.17 |
| After ICI | | | | | | |
| First | 49.37 | 20.25 | 15.19 | 21.42 | 1.01 | 93.90 |
| Second | 33.33 | 19.54 | 11.49 | 19.57 | 6.94 | 97.70 |
| Third | 56.98 | 33.72 | 8.14 | 16.25 | 8.95 | 95.35 |
| Fourth | 50.00 | 31.71 | 10.98 | 24.77 | 9.26 | 98.77 |
| Difference After ICI - Before ICI | | | | | | |
| First | 8.06 | 1.41 | 5.04 | 2.77 | 0.44 | 33.16 |
| Second | 11.02 | 6.32 | -1.73 | -9.03 | 2.26 | 24.06 |
| Third | 33.10 | 21.78 | -3.80 | -12.37 | 0.46 | 17.91 |
| Fourth | 22.93 | 19.68 | -0.30 | -3.72 | 3.12 | 29.59 |
| Obs. | 860 | 860 | 860 | 885 | 383 | 866 |

Notes. ^aYa def. [A] - Jobs included: entrepreneurs, directors, engineers and architects, business consultants, lawyers, university professors; Ya def. [B] - Jobs included: [A] less business consultants and lawyers; Ya def. [C] - Jobs included: [B] less university professors; Yp - Percentage of political experience. ^bYw def. [A] - Separate waste collection; Yw def. [B] - Completing term in office. ^cFirst: VALUE_ADDEDpc <= 13472; Second: VALUE_ADDEDpc > 13472 & VALUE_ADDEDpc <= 16447; Third: VALUE_ADDEDpc > 16447 & VALUE_ADDEDpc <= 18986; Fourth: VALUE_ADDEDpc > 18986.

Table 2. The impact of decentralization on the type and the performance of politicians (estimates of delta, 1988-1997)

| | Type | | | | |
|-------------|----------------------|----------------------|----------------------|----------------------|---------------------|
| | (1) | (2) | (3) | (4) | (5) |
| Ya def. [A] | 0.156*** (0.051) | 0.151*** (0.054) | 0.135*** (0.051) | 0.143*** (0.051) | 0.234 (0.163) |
| Ya def. [B] | 0.140*** (0.044) | 0.136*** (0.044) | 0.110** (0.046) | 0.119*** (0.046) | 0.201 (0.130) |
| Ya def. [C] | 0.040 (0.038) | 0.042 (0.038) | 0.031 (0.039) | 0.041 (0.041) | 0.019 (0.117) |
| Yp | -0.086*** (0.024) | -0.082*** (0.027) | -0.070*** (0.024) | -0.075*** (0.022) | -0.164** (0.069) |
| | Performance | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Yw def. [A] | 1.608*** (0.562) | 1.712*** (0.565) | 1.961** (0.771) | 1.568 (0.954) | 3.263** (1.435) |
| Yw def. [B] | 0.172** (0.078) | 0.171** (0.069) | 0.189*** (0.072) | 0.199*** (0.074) | 0.014 (0.092) |

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Additional controls: Model (1) - AGE, SF, POP, OVER65, MOVIPROV (CYCLE in performance model only); Model (2) - Model (1) plus ALIGNMENT, COMPETITION; Model (3) - Model (2) plus LN, FI (TL in performance model only); Model (4) - Model (3) plus interactions of political variables with VALUE_ADDEDpc; Model (5) - Model (3) plus interactions of political variables with DECENTR. Variables definition: Ya def. [A] - Jobs included: entrepreneurs, directors, engineers and architects, business consultants, lawyers, university professors; Ya def. [B] - Jobs included: [A] less business consultants and lawyers; Ya def. [C] - Jobs included: [B] less university professors; Yp - Percentage of political experience; Yw def. [A] - Percentage of separate waste collection; Yw def. [B] - Completing term in office.

Table 3. Placebo tests

| | | Type | | | Performance | |
|-----------|--------------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | | Ya | | | Yp | Yw |
| | | [A] | [B] | [C] | | [B] |
| 1988-1992 | F_Decentr1*Value_Added _{pc} | -0.067 (0.129) | -0.066 (0.099) | -0.003 (0.082) | -0.039 (0.069) | -0.067 (0.153) |
| | F_Decentr2*Value_Added _{pc} | -0.028 (0.061) | -0.046 (0.041) | -0.020 (0.040) | -0.000 (0.038) | 0.015 (0.106) |
| | | | | | | Yw |
| | | | | | | [A] |
| 1993-1994 | F_Decentr3*Value_Added _{pc} | | | | | 1.519 (1.270) |

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models (Yw de.[A] Municipal effects only). F_Decentr1*Value_Added_{pc}=0 if year<1990; f_Decentr1*Value_Added_{pc} =Value_Added_{pc} from 1990.

F_Decentr2*Value_Added_{pc}=Value_Added_{pc} starting from first municipal election held after 1989 (= 0 otherwise). F_Decentr3*Value_Added_{pc}=0 in 1993; f_Decentr3*Value_Added_{pc}=Value_Added_{pc} in 1994. Value_Added_{pc} averaged across years 1991-1992. Additional controls: AGE, SF, POP, OVER65, MOVIPROV, ALIGNMENT, COMPETITION (CYCLE in performance model only). Variables definition: Ya def. [A] - Jobs included: entrepreneurs, directors, engineers and architects, business consultants, lawyers, university professors; Ya def. [B] - Jobs included: [A] less business consultants and lawyers; Ya def. [C] - Jobs included: [B] less university professors; Yp - Percentage of political experience; Yw def. [A] - Separate waste collection; Yw def. [B] - Completing term in office.

Table 4. The impact of decentralization on the type and the performance of politicians (estimates of delta, 1995-2002)

| | Type | | | | |
|-------------|--------------------|--------------------|--------------------|-------------------|--------------------|
| | (1) | (2) | (3) | (4) | (5) |
| Ya def. [A] | 0.005 (0.040) | 0.002 (0.041) | 0.001 (0.041) | -0.013 (0.038) | 0.067 (0.148) |
| Ya def. [B] | 0.043 (0.047) | 0.039 (0.046) | 0.037 (0.044) | 0.021 (0.043) | 0.162 (0.134) |
| Ya def. [C] | 0.086** (0.039) | 0.081** (0.036) | 0.080** (0.035) | 0.057* (0.033) | 0.300** (0.117) |
| Yp | -0.031* (0.017) | -0.031* (0.017) | -0.031* (0.018) | -0.023 (0.016) | -0.018 (0.062) |
| | Performance | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Yw def. [A] | 1.494** (0.760) | 1.421** (0.648) | 1.424** (0.666) | 0.978* (0.565) | 1.091 (1.331) |
| Yw def. [B] | 0.003 (0.033) | 0.003 (0.033) | -0.011 (0.025) | 0.005 (0.027) | -0.098 (0.070) |

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Additional controls: Model (1) - AGE, SF, POP, OVER65, MOVIPROV (CYCLE in performance model only). Model (2) - Model (1) plus ALIGNMENT, COMPETITION; Model (3) - Model (2) plus LN, FI (TL in performance model only); Model (4) - Model (3) plus interactions of political variables with VALUE_ADDEDpc; Model (5) - Model (3) plus interactions of political variables with DECENTR. Variables definition: Ya def. [A] - Jobs included: entrepreneurs, directors, engineers and architects, business consultants, lawyers, university professors; Ya def. [B] - Jobs included: [A] less business consultants and lawyers; Ya def. [C] - Jobs included: [B] less university professors; Yp - Percentage of political experience; Yw def. [A] - Separate waste collection; Yw def. [B] - Completing term in office.

Table 5. The role of spending for electoral campaigns (estimates of delta)

| | Type | | | Yp | Performance | |
|-----------|--------------------|--------------------|-------------------|----------------------|--------------------|--------------------|
| | Ya | | [C] | | Yw | |
| | [A] | [B] | | | [A] | [B] |
| 1988-1997 | 0.118** (0.053) | 0.093** (0.041) | 0.038 (0.031) | -0.070*** (0.025) | 2.122** (0.846) | 0.166** (0.080) |
| 1995-2002 | 0.021 (0.040) | 0.067 (0.044) | 0.070* (0.036) | -0.034** (0.016) | 1.682** (0.732) | -0.000 (0.025) |

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Additional controls: AGE, SF, POP, OVER65, MOVIPROV, ALIGNMENT, COMPETITION, LN, FI (CYCLE, TL in performance model only). Variables definition: Ya def. [A] - Jobs included: entrepreneurs, directors, engineers and architects, business consultants, lawyers, university professors; Ya def. [B] - Jobs included: [A] less business consultants and lawyers; Ya def. [C] - Jobs included: [B] less university professors; Yp - Percentage of political experience; Yw def. [A] - Separate waste collection; Yw def. [B] - Completing term in office.

Table 6. The role of social capital (estimates of delta)

| | | Type | | | | Performance | |
|-----------|--------------|---------------------|--------------------|--------------------|---------------------|------------------|---------------------|
| | | Ya | | | Yp | Yw | |
| | | [A] | [B] | [C] | | [A] | [B] |
| 1988-1997 | Centre-North | 0.141*** (0.046) | 0.109** (0.047) | 0.015 (0.037) | -0.066** (0.027) | 0.723 (1.055) | 0.189** (0.083) |
| | South | -0.019 (0.186) | 0.129 (0.124) | 0.111 (0.104) | -0.099** (0.045) | 1.305 (0.946) | 0.361*** (0.124) |
| 1995-2002 | Centre-North | 0.011 (0.045) | 0.027 (0.047) | 0.075** (0.034) | -0.029* (0.018) | 0.965 (0.631) | -0.001 (0.032) |
| | South | -0.001 (0.095) | 0.024 (0.090) | 0.046 (0.066) | -0.036 (0.045) | 1.074 (1.197) | -0.046 (0.065) |

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Additional controls: AGE, SF, POP, OVER65, MOVIPROV, ALIGNMENT, COMPETITION, LN, FI (CYCLE, TL in performance model only) Variables definition: Ya def. [A] - Jobs included: entrepreneurs, directors, engineers and architects, business consultants, lawyers, university professors; Ya def. [B] - Jobs included: [A] less business consultants and lawyers; Ya def. [C] - Jobs included: [B] less university professors; Yp - Percentage of political experience; Yw def. [A] - Separate waste collection; Yw def. [B] - Completing term in office.

Table 7. Mutually exclusive definition of Y_a and Y_p (estimates of delta)

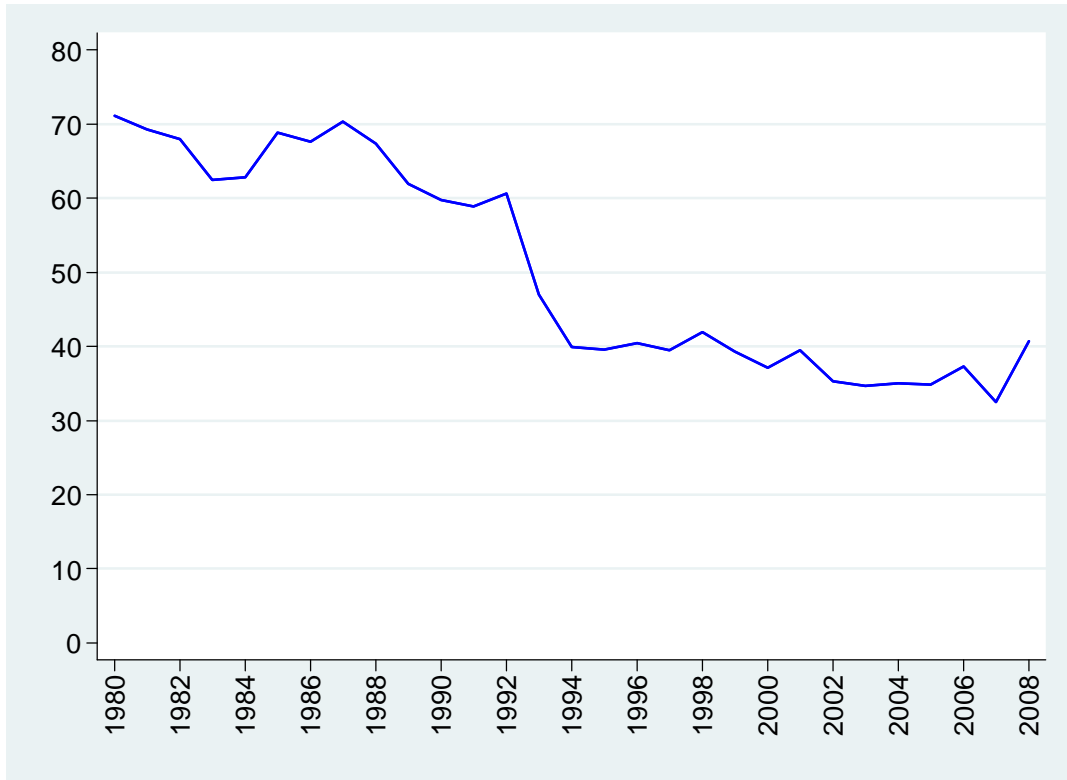
| | | Ya definitions | | |
|-----------|-------|---------------------|---------------------|--------------------|
| | | [A] | [B] | [C] |
| 1988-1997 | Y_a | 0.136*** (0.051) | 0.121*** (0.045) | 0.034 (0.037) |
| | Y_p | -0.020 (0.034) | -0.017 (0.037) | -0.017 (0.037) |
| 1995-2002 | Y_a | 0.018 (0.040) | 0.045 (0.043) | 0.082** (0.034) |
| | Y_p | 0.018 (0.020) | 0.010 (0.024) | 0.004 (0.024) |

Notes. Two-ways clustered SE in parentheses. Sig. Lev.: *** 1%, ** 5%, * 10%. Municipal and year FE included in all models. Additional controls: AGE, SF, POP, OVER65, MOVIPROV, ALIGNMENT, COMPETITION, LN, FI. Variables definition: Y_a def. [A] - Jobs included: entrepreneurs, directors, engineers and architects, business consultants, lawyers, university professors; Y_a def. [B] - Jobs included: [A] less business consultants and lawyers; Y_a def. [C] - Jobs included: [B] less university professors; Y_p - Percentage of political experience.

Appendix Table A.1. Descriptive statistics (1988-1997)

| Variable | Obs | Mean | Std. Dev. | Min | Max | Source |
|---------------------------|-----|--------|-----------|--------|---------|---|
| Ya [A] | 860 | 0.360 | 0.480 | 0 | 1 | Italian Ministry of Domestic Affairs, |
| Ya [B] | 860 | 0.188 | 0.391 | 0 | 1 | "Anagrafe degli amministratori locali" |
| Ya [C] | 860 | 0.115 | 0.319 | 0 | 1 | |
| Yp | 885 | 0.238 | 0.199 | 0 | 0.917 | Italian Ministry of Domestic Affairs, "Anagrafe degli amministratori locali"; OPENOPOLIS; historical archive "Corriere della Sera" and "La Repubblica"; Italian and European Parliament |
| Yw [A] | 383 | 6.371 | 6.502 | 0 | 33.740 | Legambiente |
| Yw [B] | 866 | 0.804 | 0.397 | 0 | 1 | Italian Ministry of Domestic Affairs, "Anagrafe degli amministratori locali" |
| Value_Added _{pe} | 890 | 16298 | 3731 | 8851 | 26041 | Istituto Tagliacarne |
| SF | 885 | 0.045 | 0.208 | 0 | 1 | Italian Ministry of Domestic Affairs, "Anagrafe degli amministratori locali" |
| AGE | 885 | 50 | 9 | 29 | 81 | Italian Ministry of Domestic Affairs, "Anagrafe degli amministratori locali" |
| POP | 890 | 174380 | 345264 | 15008 | 2800000 | ISTAT |
| OVER65 | 890 | 17.005 | 3.811 | 7.690 | 26.560 | |
| MOVIPROV | 890 | 7.026 | 1.421 | 4.050 | 13.149 | Italian Chambers of Commerce |
| ALIGNMENT | 885 | 0.702 | 0.458 | 0 | 1 | Italian Ministry of Domestic Affairs, "Anagrafe degli amministratori locali"; OPENOPOLIS; historical archive "Corriere della Sera" and "La Repubblica" |
| COMPETITION | 890 | 2145.2 | 390.1 | 1482.9 | 3657.2 | Italian Ministry of Domestic Affairs, "archivio storico delle elezioni"; ITANES |
| LN | 885 | 0.044 | 0.205 | 0 | 1 | Italian Ministry of Domestic Affairs, |
| FI | 885 | 0.028 | 0.166 | 0 | 1 | "Anagrafe degli amministratori locali"; OPENOPOLIS; historical archive "Corriere della Sera" and "La Repubblica" |
| TL | 890 | 0.018 | 0.133 | 0 | 1 | Italian Ministry of Domestic Affairs, |
| CYCLE | 889 | 2.781 | 1.385 | 1 | 5 | "Anagrafe degli amministratori locali" |

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



Source: ISTAT, Conto economico delle amministrazioni comunali

Appendix [For Online Publication]

Derivation of r_1^{j*}

Rewriting Eq.(5) fully,

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

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

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

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

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

Recall from Eq. (3) 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.2, this clearly holds for both $z = E(z^a)$ and $z = E(z^p)$.

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

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

that solving for r_1^j , gives

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

QED.

Proof of Proposition 1

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

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

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

$$R' = \frac{E(z^k)}{\left(1 - \frac{\lambda}{2} + \frac{\psi \lambda}{\tau} [E(z^k) + s 2(E(z^{-k}) - E(z^k))]\right)} \geq \bar{R} \quad (9)$$

Notice from (6), that (9) also implies $r_1^{k*} > 0$. Second, the voter must also prefer to play his equilibrium strategy to any other possible alternative. The strategy of always failing the weakest expected candidate k at the elections clearly strictly dominates the strategy of having him always confirmed (or any random choice between always defeating and always re-electing him). Under this alternative best strategy, the expected welfare of the consumer is then:

$$\tilde{U} = E(z^k) - \bar{R} + \lambda [E(z^o) - \bar{R}] \quad (10)$$

Computing the expected welfare of the consumer at the proposed equilibrium is instead:

$$U^* = \frac{\lambda \bar{R} \psi}{\tau} E(z^o) + \lambda \left[\frac{s}{2} (E(z^k) + E(z^k | z^k \geq E(z^k))) + (1-s) ((1-q^k) E(z^{-k}) + q^k E(z^k | z^k \geq E(z^{-k}))) \right] - \lambda \bar{R} \quad (11)$$

where $s = \pi$ if $k = a$ and $s = (1 - \pi)$ otherwise, and $q^k = \frac{1}{2} + \frac{\psi}{\tau} [E(z^k) - E(z^{-k})]$. It follows that a *sufficient* condition for $U^* \geq \tilde{U}$ is:

$$\bar{R} \geq \frac{E(z^k)}{\left(\frac{\lambda \psi}{\tau} E(z^o) + 1\right)} = R'' \quad (12)$$

Combining (9) and (12), we see that $R' > R''$ if $m = \frac{1}{2s[E(z^{-k}) - E(z^k)]} > \frac{\psi}{\tau}$, which is implied by A.2. Hence, $R' > R''$. As stated in Proposition 1, for $R' > \bar{R} > R''$ we then get an equilibrium. To prove that this equilibrium is also unique, compute $r_1^j(r_1^{je}) = \arg \max E(R^j)$ for an arbitrary value of r_1^{je} and note that $r_1^j(r_1^{je}) = r_1^{je}$ only for $r_1^{j*} = 1 - \frac{\lambda \bar{R} \psi}{\tau} \left[\frac{E(z^o)}{E(z^j)} \right]$. QED.

Proof of Proposition 2

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

- for a -type

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

Solving:

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

which is positive if

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

This holds if

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

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

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

Consider now

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

which is positive if

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

This holds if

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

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

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

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

- For p -type:

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

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

and

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

Finally:

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

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

QED.

Proof of Proposition 3

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

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

$$\begin{aligned} U^p = & \frac{\lambda \bar{R} \psi}{\tau} E(z^o) - \lambda \bar{R} + \lambda \left((1 - \pi) \left(\frac{1}{2} (E(z^p) + \right. \right. \\ & \left. \left. + E(z^p | z^p \geq E(z^p))) + \lambda \pi \left(\left(\frac{1}{2} - \frac{\psi}{\tau} h \right) E(z^p | z^p \geq E(z^a)) + \right. \right. \\ & \left. \left. + \left(\frac{1}{2} + \frac{\psi}{\tau} h \right) E(z^a) \right) \right) \end{aligned} \quad (14)$$

where $h = E(z^a) - E(z^p)$. Invoking (3), and solving the integrals we obtain:

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

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

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

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

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

Substituting in (13) and (14) and simplifying:

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

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

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

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

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

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

The first two terms are strictly positive; the third is generally uncertain. But going through element by element, it is easy to check that the third term is dominated by the

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

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

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

Derivation of π^*

Consider again the expected utility of the two types of politicians

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

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

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

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

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

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

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

Summing and substituting we obtain:

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

which can be rewritten as:

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

Let $A + B = x$; $\lambda\bar{R} + (a + p)(1 - k) = b$; $2(a - p)^2k(1 + k) = c$, leading to $x^2 - bx + c = 0$. Solving, the two roots are $\frac{1}{2}b + \frac{1}{2}\sqrt{b^2 - 4c}$; $\frac{1}{2}b - \frac{1}{2}\sqrt{b^2 - 4c}$. The equation admits real solutions if $b^2 \geq 4c$. This is certainly the case for $|a - p| \leq \frac{b}{2\sqrt{2(k+k^2)}} = Q$. Note that Q is decreasing in k and that for $k \rightarrow 1$, $Q \rightarrow \frac{\tau}{4\psi}$. As $k < 1$, $|a - p| \leq Q$ is then a very mild condition, that is already implied by A.2. Note further that to make economic sense $|\frac{(a-p)(1+k)}{2(A+B)}| \leq \frac{1}{2}$; this would certainly be violated by the negative root for $a \rightarrow p$. Thus, the only economic significant solution is represented by the positive root. Under the positive root, $A + P = \frac{n(b + \sqrt{b^2 - 4c})}{2\bar{w}(1 + \lambda)}$. Substituting, this gives equation (7) into the main text. QED.

Proof of Proposition 4

Differentiating the numerator and the denominator of (7) for t and imposing $\tau = 1 - \alpha t$, we get:

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

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

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

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

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

It follows:

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

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