FISCAL FEDERALISM AND POLITICAL SELECTION: EVIDENCE FROM ITALY

MASSIMO BORDIGNON, MATTEO GAMALERIO, GILBERTO TURATI

FISCAL FEDERALISM AND POLITICAL SELECTION: EVIDENCE FROM ITALY

Massimo Bordignon
Catholic University, Milan

Matteo Gamalerio
LSE

Gilberto Turati
University of Turin

June 2012
Very preliminary

1 Introduction

As is well known, the empirical case for fiscal decentralization is in general terms rather weak. While the theoretical literature typically emphasizes the potentially large efficiency gains deriving from decentralization (Lockwood, 2006 for a recent survey), the empirical evidence is by far 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). One argument that is often put forward in the literature as a potential explanation for these contrasting results concerns the mismatch at the local level between own revenues and expenditure: the efficiency gains associated to decentralization are in general much more likely to be realized when the degree of "vertical imbalance" (the percentage of transfers on local government total revenues) is low than when it is high (see Eyraud and Lusinyan, 2011, for recent cross-country strong supportive evidence). But why this is the case has never been made very clear in the literature, although several informal suggestions are routinely advanced in the debate. In this paper, using the Italian case as a testing ground for the analysis, we focus on a possible channel of causation, the relationship between the "quality" of decentralization – that is, whether it involves or not high levels of vertical imbalances – and the "quality" of (local) politicians.

Our main idea, more formally discussed in Section 2, is the following. There are several reasons why one should expect the characteristics of local politicians to be endogenous to the "quality" of decentralization. For example, in a decentralized setting where most of the resources still come from the center, the main task of the 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, etc.) than those of a good administrator of local matters. And given the choice,

voters of communities with high degrees of vertical imbalance would rationally prefer the former type of politician to the latter. More generally, there can also be a self-selection effect. Candidates of different abilities may decide to enter in the local political arena in communities with different level of vertical fiscal imbalances. Finally, and not in conflict with the previous explanations, the political parties themselves may prefer to propose different types of candidates in contexts with different levels of autonomous resources.

The Italian experience is particularly suitable to address these issues. For, as explained in more details in Section 3, the 1993 reforms changed both the electoral system of municipalities – moving from a pure proportional- parliamentary system to a quasi-presidential one, involving the direct elections of mayors and of a municipal legislature majority directly linked to the elected mayor – and the funding of municipalities, introducing a new tax resource, the municipal property tax, that dramatically improved the financial system 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 communities, the effect of the tax reform was to make the municipalities almost entirely financially independent from the center; in the poorest ones, in spite of the reform, municipalities kept receiving most resources in terms of transfers from the center. We posit that this should have affected 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.

In order to test this hypothesis, we then collect an extensive data set on the personal characteristics of the mayors of the main Italian cities (those that are Capoluogo di Provincia, literally Chief provincial towns, approximately 90 cities in the Italian Ordinary Statute Regions), as well as on other economic and political characteristics of the cities. Focusing only on the mayor (and not on the general features of the local political class), aside of data availability, is justified by the paramount role in municipal policy attributed to this figure by the 1993 reform. Building on an extensive literature in political economics (e.g., Nannicini and Gagliarducci, 2011; Nannicini and Galasso, 2011), we then construct proxy measures for different "types" of mayors as well as measures of "ex-post" quality of policy, looking at a specific output indicator related to one of the main services offered by the municipalities (namely, waste collection).

Results are strongly supportive of our working hypothesis. In the municipalities where the electoral reform was accompanied by a large increase in autonomous resources, the ex-ante quality of politicians changed dramatically in the aftermath of the reform. A much larger percentage of elected mayors came from top professions in the private sector. There is also some evidence that in these municipalities the higher quality of the political local class was also reflected in a higher quality of policy ex-post, although we can measure it only very imperfectly. On the contrary, we observe no effect or very little effect of the electoral reform on the poorer communities both in terms of the ex-ante quality of politicians and ex-post quality of policy. We provide a number of robustness test for these results, checking whether they were not driven by other factors that the literature typically associates with a better selection of politi-

cians, and by the peculiar period under study. For instance, our results hold even controlling for the degree of competitiveness in the electoral competition at the local level (e.g., Nannicini and Galasso, 2011) and for the endowment of "social capital" at municipal level (Guiso, Sapienza and Zingales, 2010), and for a different time period related to a different change in the financing of Italian municipalities (namely, the introduction of a surcharge on the Personal Income Tax).

Our results have strong implications for the debate on fiscal federalism, both in Italy as elsewhere. Clearly, not all recipes are adaptable to all circumstances. Decentralization may be a good idea, but it requires appropriate conditions, in particular a sufficient degree of local financial autonomy, in order to work.

Our work is related to several different lines of research. 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 work that attempts to do so is a recent work by Brollo et al. (2010) on Brazilian municipalities. The authors suggest that higher transfers might lead to a poorer quality of politicians, but in their case the result is mechanical due to the larger resources that with higher transfers a bad politician can expropriate when elected. In our case, it is more a matter of self-selection of politicians with different political skills in different financial contexts. Our work is also related to the recent strand of research in political economics that has increasingly focused on the effects of the different political institutions on the selection of good politicians (e.g., Besley, 2006). While most effort in this recent literature has been so far devoted to address the relationship between compensation and the quality of politicians , it is clear that the insight is much more general and could be applied to other types of institutions.

The remainder of the work is organized as follows. Section 2 presents a theoretical model of self-selection of local politicians under different financial conditions that captures our main idea. Section 3 discusses 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, discussing also some robustness tests. Section 6 concludes.

2 The Theoretical Model

To fix ideas and gain some insights that help us in examining our data, it is useful to set up a simple model. The framework is basically the same of Brollo et al. (2010), but we extend it to different types of politicians with specialized skills. The model is a compromise between the need to make a general point and the need to offer insights on the effect of a decentralization reform, similar to the one implemented in Italy in the 90's, on the selection of politicians. It could easily be generalized, but we deliberately prefer to focus here on simplicity rather than generality, choosing functional forms and imposing simplifying assumptions that

allow us to derive analytical solutions.

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 in the second period; the world ends with the end of this second period[1]. All politicians are "bad"; they only care about collecting as much rents as possible from office. We let R_t indicate the rents taken by an incumbent in period t. Politicians come of two types, $j = a, p$; in a sense to be made more precise below, type a is on average better in organizing local services (he has more "administrative" skills), type p is on average better in raising money from the center (he has more "political" skills). Our basic point in this paper is indeed that these are quite different skills, require a different type of background and specialization, and are therefore typically distributed differently across the population of (potential) politicians.

Local taxes are fixed, or with little variation allowed (an assumption that certainly fits the Italian case), so that the voter is basically only interested in the quality/quantity of local public services, that we capture here with a single public good g_t. The utility of the voter is then just:

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

where $0 < \lambda < 1$ is the discount rate and expectations in (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 that g_1 is given by:

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

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

Note that (2) implies that the amount of public good produced locally not only depends on the total revenues accruing to the municipality, $\alpha t + \tau(\alpha)$, but also on the ability of the politician j to use these resources, captured here by the couple $\left(\theta^j, \delta^j\right)$. Intuitively, a given amount of local resources, αt, can generate a higher level of public good production if they are managed by a politician j that

[1] The model can be easily generalized to multiple periods; see for instance chapter in Persson and Tabellini, 2000.

has a higher level of administrative skills than another politician k, $\theta^j > \theta^k$. By the same token, a given level of transfer, $\tau(\alpha)$ can generate a higher level of public good, if it is managed by a politician with higher political skills, δ^j, as this politician may be more able to get extra resources from the center, or equivalently, as he may be more able to convince the center to directly finance some components of local expenditure. Notice that we also implicitly assume in (2) that the politician can divert to his advantage not only $\alpha t + \tau(\alpha)$, but also the extra resources that he himself generates. This is reasonable; a politician with high political skills may cash some of the extra transfer he brings home; a politician with high tecnical 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, note that in (2), for analytical convenience, we do not allow politicians to take different rents from the different sources of financing; the same rate r_1^j applies to both sources.

Both θ^j and δ^j follow an independent uniform distribution function with density ψ and average $\overline{\theta}^j > 0$, $\overline{\delta}^j > 0$, respectively. In keeping with the discussion above, we also assume $\overline{\theta}^a > \overline{\delta}^a, \overline{\delta}^p > \overline{\theta}^p$, $\overline{\theta}^a > \overline{\theta}^p$, $\overline{\delta}^a < \overline{\delta}^p$. This captures the idea that a types are on average "better" in producing local services out of local resources, while p types are on average "better" in producing local services out of the resources coming from the center. This formulation, while simple, is quite general, and could be easily amended to introduce other realistic features in the model[2].

Let $z^j = \alpha t \theta^j + \tau(\alpha)\delta^j$ be the total municipal income generated by an incumbent of type j and $f(z^j)$ 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 uniformely distributed random variables, with density $\frac{\psi}{\alpha t}$ and $\frac{\psi}{\tau(\alpha)}$, respectively. The extremes of the two variables are: $\underline{y}^j = \tau(\alpha)(-\frac{1}{2\psi} + \overline{\delta}^j)$; $\overline{y}^j = \tau(\alpha)(\frac{1}{2\psi} + \overline{\delta}^j)$, $\underline{x}^j = \alpha t(-\frac{1}{2\psi} + \overline{\theta}^j)$, $\overline{x}^j = \alpha t(\frac{1}{2\psi} + \overline{\theta}^j)$. 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 $\tau(\overline{\alpha}) > \overline{\alpha} t$;

As a justification, one can notice that $\tau(\alpha) > \alpha t$ indeed characterized the situation of all municipalities in pre-reform Italy, and in most of them even

[2] For instance, one may imagine that a citizen cares more for the politician diversion of funds coming from her pocket than from a diversion of the transfers, as the latter are paid by someone else. This could be captured by writing the utility function of the representative voter (in the first period) more generally as

$U_1 = g_1 + (1-t)\alpha - \lambda t = \left[\alpha t \theta^j + \tau \beta \delta^j\right](1 - r_1^j) + (1-t)\alpha - \lambda(\alpha)t$,

where $(1-t)\alpha$ is private consumption net of local taxation and $\lambda(\alpha) > 0$ is some measure of the deadweight loss of taxation. With a variable local tax rate, this immediately implies a higher utility loss for the consumer out of local resources, which in turn, should lead the voter to punish more severely politicians that use local funds to accumulate rents. In our case, however, for simplicity we keep t fixed, so that we can neglet the extra terms.

after the reform (see next section). Finally, let $k = \frac{\psi^2}{\alpha t \tau(\alpha)}$. By the convolution theorem, under A.1, $f(z^j)$ can be computed as:

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

Finally, we also impose an exogeneous 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 $j, k = a, p$, so simplifying the computations below.

2.1 The equilibrium

The game enfolds as follows. In the first period, the incumbent politician j chooses r_1^j knowing $f(z^j)$, but before observing the realization of z^j. This assumption is introduced to simplify the analysis (it avoids signalling problems), but it may also have some bearing with reality. 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 is only known after he has tried his hands in govern. Notice that this implies that all politicians of type j, as they are ex-ante identical, make the same choice of r_1^j in period 1. Notice further that ex post, once θ^j and δ^j and therefore z^j are realized, as these are personal characteristics of the incumbent, their realization in the first period carry over to the second. This implies that the voter has an incentive to re-elect an incumbent with better realized characteristics than the expected z^o of an opponent. Furthermore, as both θ^j and δ^j are realized before the elections (the candidate proves his worth by ruling), g_1^j is also determined and the voter can then condition her vote on g_1^j. In the second period, whoever is in charge, say type k, chooses again some rents, r_2^k , and g_2^k is realized. The game then ends.

To solve the model, we work backwards. In period 2, as there is no future ahead, whoever is in charge takes maximal rents, $R_2^k = \overline{r} z^k$, where $\overline{r} < 1$ is some maximal rent rate. For analytic simplicity, we assume here that maximal rents an incumbent can take in each period take some fix values, independently on j and z^j, i.e. $R_s^k = \overline{R} > 0$ for $k = a, p$ and $s = 1, 2$ (the case with $R_s^k = \overline{r} z^k$ is worked out in the Appendix). In the second period, the utility of the voter is then $z^k - \overline{R}$, so that the voter is however interested to re-elect or elect 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, knows $f(z^j)$ and knows the type of the incumbent politician, that is, she knows whether $j = a$ or $j = p$, 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 $\overline{R}$ in the first period too). We then look for an equilibrium where the voters use these expectations to discriminate between high quality / low quality incumbents. Let r_1^{je} be the rate of rents that the voter expects a politician of type j to take in period 1. Upon observing g_1^j, the expected value of z^j for the voter is then just:

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

Intuitively, the best strategy of the voter is then to vote for the incumbent if $E(z^j|g_1) \geq E(z^o)$ and vote for the opponent otherwise. At the equilibrium, the incumbent knows the voter's optimal strategy when setting r_1^j. Ex ante, he can then compute the probability of being relected as a function of r_1^j, that is, $prob(z^j \frac{(1-r_1^j)}{(1-r_1^{je})} \geq E(\ z^o))$. We suppose that in period 1, when he sets up r_1^j, the incumbent does not know the type of the opponent he is going to face at the elections; this makes sense intuitively (in Italy, the opponent in local elections is usually unknown until few months before the elections). In particular, we suppose that there is a fraction π of type a politicians in the population (to be endogeneized below) and that the opponent is selected randomly from the poulation just before the elections.

Under these assumptions, the expected rents of an incumbent of type j over the two periods can be written as:

$$\begin{aligned} E(R^j) \quad = \quad & r_1^j E(z^j) + \lambda \overline{R} \pi \left[1 - prob\left(z^j \leq E(z^a) \frac{(1 - r_1^{je})}{(1 - r_1^j)} \right) \right] + \\ & + \lambda \overline{R} (1 - \pi) \left[1 - prob\left(z^j \leq E(z^p) \frac{(1 - r_1^{je})}{(1 - r_1^j)} \right) \right] \end{aligned} \tag{5}$$

Invoking (3) and A.2, the probability of being re-elected for the incumbent j as a function of r_1^j, let us call it $h^j(r_1^j)$, around the equilibrium level of rents [3] can be computed as:

[3] At the equilibrium, $r_1^j = r_1^{je}$ must hold. This implies that in order to verify that $r_1^j = r_1^{je}$ is indeed a (local) maximum, one must consider how the probability of being re-elected for incumbent j would change for small variations in r_1^j around the point $r_1^j = r_1^{je}$. A.2 implies that $\overline{x}^j + \underline{y}^j \leq E(z^j), E(z^k) \leq \underline{x}^j + \overline{y}^j$ so that the relevant part of the distribution function $f(z^j)$ to do this computation is its "flat" component. This leads to (7).

This implies that the relevant portion of the distribution function $f(z^j)$ to be used to verify that $r_1^j = r_1^{je}$ is indeed a local maximun is

$$h^j(r_1^j) = \frac{1}{2} - \frac{\psi}{\tau}\pi\left[E(z^a)\frac{(1-r_1^{je})}{(1-r_1^j)}) - E(z^j)\right] + \quad (6)$$
$$-\frac{\psi}{\tau}(1-\pi)\left[E(z^p)\frac{(1-r_1^{je})}{(1-r_1^j)}) - E(z^j)\right]$$

where we suppress the argument in $\tau(.)$ to simplify the notation. Note that $h^{j\prime}(r_1^j) < 0$ and $h^{j\prime\prime}(r_1^j) < 0$, so that the function $E(R^j)$ is strictly (locally) concave. Substituting from (6) in (5), deriving $E(R^j)$ with respect to r_1^j, setting $\partial E(R^j)/\partial\, r_1^j = 0$, and imposing the equilibrium condition $r_1^j = r_1^{je}$, we get the equilibrium rents rate, r_1^{j*}:

$$r_1^{j*} = r_1^{je} = 1 - \frac{\lambda\overline{R}\psi}{\tau}\left[\frac{E(z^o)}{E(z^j)}\right] \quad (7)$$

where $E(z^o) = \pi E(z^a) + (1-\pi)E(z^p)$. At the equilibrium, expected rents in the first period for the $j's$ incumbent are then $r_1^{j*}E(z^j) = E(z^j) - \frac{\lambda\overline{R}\psi}{\tau}E(z^o)$. Thus, expected rents are increasing in $E(z^j)$; more ex ante efficient types get higher rents at the equilibrium. They are instead decreasing in $\lambda\overline{R}$ (a larger $\lambda\overline{R}$ means that future 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^{je}). 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 $h^{jk} = \frac{1}{2} + \frac{\psi}{\tau}\left[E(z^j) - E(z^k)\right]$, $j,k = a,p$, $j \neq k$, if he meets a candidate of a different type[4]. So, at the proposed equilibrium, more efficient types also expect to be re-elected with higher probability.

The strategies of the voter at the equilibrium are also straightfoward: when two opponents of the same type, say j, compete at the elections, the voter re-elects the incumbent iff $g_1 \geq \frac{\lambda\overline{R}\psi}{\tau}E(z^o)$; when an incumbent of type j compete with a candidate of type $k \neq j$, the voter re-elect the incumbent iff $g_1 \geq \frac{\lambda\overline{R}\psi}{\tau}E(z^o)\frac{E(z^k)}{E(z^j)}$. The intuition is simple. If a candidate of type j is known to be more efficient in expected terms that a candidate of type k the voter is willing to re-elect j even if he observes a g_1 smaller than $g_1^* = \frac{\delta\overline{R}\psi}{\tau}E(z^o)$, provided that is larger than $k's$ expected value, $g_1^*\frac{E(z^k)}{E(z^j)}$.

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

[4]Note that by invoking A.2), $0 < h^{jk} < 1$.

elections. As the Appendix proves, these conditions simply translate in upper and lower limits for $\overline{R}$, $\overline{R} \subset \{R', R''\}$. We can then conclude:

Proposition 1 *Assume $R' > \overline{R} > R''$. Then there exists an equilibrium where the voter sets up a threshold for g_1, $g_1^* = \frac{\lambda \overline{R} \psi}{\tau} E(z^o) \frac{E(z^k)}{E(z^j)}$ such that she re-elects the incumbent j, if $g_1 \geq g_1^*$, and she elects the opponent k otherwise (where $j, k = a, p$). At this equilibrium, an incumbent j sets first period rents at the rate $r_1^{j*} = 1 - \frac{\lambda \overline{R} \psi}{\tau} \left[\frac{E(z^o)}{E(z^j)} \right]$, where $0 < r_1^{j*} < 1$. $E(z^j) \geq E(z^k)$ $(E(z^j) \leq E(z^k))$ implies that incumbent j 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, $j \neq k$.*

For future references, it might also be interesting to ask whether the consumer is better off or worse off in expected terms when the incumbent is of the most efficient type. In the first period, the consumer is of course indifferent; she always receives in expected terms $g_1 = \frac{\lambda \overline{R} \psi}{\tau} E(z^o)$ regardless of the type. In the second period, however, the welfare of the voter depends on the type of the incumbent, as the support for $E(z^j | z^j \geq E(z^k))$ depends on j, where $j, k = a, p$. What about consumer total expected utility, that is, taking into account second period expected consumption?

Lemma 2 *Let U^j be the expected utility of the consumer at the equilibrium where she faces an incumbent of type j in the first period. Then: i) sign$(U^a - U^p)$=sign$(E(z^a) - E(z^p))$; ii) if $(U^a - U^p) > 0$ (< 0), $(U^a - U^p)$ decreases (increases) as π increases.*

The consumer is then certainly better off in expected terms if he is governed by the most efficient incumbent in the first period. However, the extent of this expected gain depends on the share of the different types in the population of politicians. To understand the intuition, suppose $E(z^a) > E(z^p)$ and consider the extreme case $\pi \to 0$. Here the benefit for the consumer of having an efficient incumbent a is maximal, because the opponents are p-types only, and therefore the possibility of electing a more efficient type in the second period only occurs when the incumbent is of an a-type. As π increases, this advantage diminishes. Reversing sign and repeating the argument, the same also occurs when $E(z^a) \leq E(z^p)$.

2.2 Comparative statics

The simple structure of the model makes it easy to investigate the effects of changes of the parameters on the equilibrium. In particular, we are interested in studying the effects of changes in t and τ in municipalities with a different tax base α, as this is our case study, the Italian decentralization refom of the 90's. As explained in more details in the next session, with the introduction of the property tax, the central government also reduced the transfers to compensate for the new source of revenues, so that at the statutory level of the new property

tax rate, each municipality had exactly the same resources both after and before the reform. In terms of our model, the simplest way to capture this feature of the reform is by normalizing municipal revenues to unity and writing $\tau(\alpha) = 1 - \alpha t$. 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 $\alpha^* = \frac{\overline{\delta}^p - \overline{\delta}^a}{t(\overline{\theta}^a - \overline{\theta}^p + \overline{\delta}^p - \overline{\delta}^a)}$, provided that $\alpha^* \subset [\overline{\alpha}, \underline{\alpha}]$. For concreteness, we assume the latter 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 the reform in the context of our model? Consider first the expected welfare of the two types of politicians[5].

Proposition 3 *(i)* $\partial(E(R^a) - E(R^p))/\partial t > 0$; $\partial^2(E(R^a) - E(R^p))/\partial\alpha\partial t > 0$; *(ii)* $\partial E(R^p)/\partial t < 0$; $\partial^2 E(R^p)/\partial\alpha\partial t < 0$. *(iii) Suppose* $\overline{\theta}^a \geq \theta^* > 0$; *then* $\partial E(R^a)/\partial t > 0$. *Suppose* $\overline{\theta}^a \geq \theta^{**} > \theta^*$, *then* $\partial^2 E(R^a)/\partial\alpha\partial t > 0$.

To understand the proposition, 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 above assumptions on $\overline{\theta}^j$ and $\overline{\delta}^j$ this effect is certainly positive for type a, as $(\partial E(z^a)/\partial t - \alpha\partial E(z^a)/\partial\tau)dt > 0$, and certainly negative for type p, as $(\partial E(z^p)/\partial t - \alpha\partial E(z^p)/\partial\tau)dt < 0$. Notice that this also implies that a politician of type a is also more likely to be re-elected in the second period (when meeting an opponent of a different type) as $(E(z^a) - E(z^p))$ increases. But there is also an indirect effect; the change in t and therefore in τ increases the density around the equilibrium; $-\frac{\psi}{\tau^2}d\tau/dt > 0$.This indirect effect is *negative* for both types in the first period, as it reduces the expected rents in the first period (see the Appendix for a formal proof); but it is also negative (positive) in the second period for type p (type a) as it reduces (increases) the probability of being re-elected. Thus, the p type of 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 $\overline{\theta}^a$ being larger of a given threshold, θ^*. Notice also that $\partial^2 E(R^p)/\partial\alpha\partial t < 0$, meaning that the loss induced of the reform on the p-type is higher in municipalities with a higher tax base, in spite of the fact that the reform perfectly compensated (at the statutory tax base) for any difference in the tax base across municipalities. Conversely, $\partial^2 E(R^a)/\partial\alpha\partial t > 0$ provided that $\overline{\theta}^a$ is again larger of a given, higher, threshold, $\theta^{**} > \theta^*$. Thus, while the p-type is certainly made worse off by the reform, the effects on the a-type are in general uncertain. But the important point, also stated in Proposition 2, is that regardless of its effects on the absolute level

[5] In proposition 2, as in the propositions to follow, when we differentiate for t we take into account the dependence of τ on t, that is $\tau(\alpha) = 1 - \alpha t$.

of welfare, 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? Her utility certainly increases in the first period, as first period rents for both types fall after the reform. But in the second period the effects are complex as they depend on the absolute value of the difference between $E(z^a) - E(z^p)$, the identity of the incumbent politician and the relative size of two types of politicians in the population. Thus, it is in general impossible to sign unambiguosly the effect of the reform on the welfare of the consumer. However, as we conjecture in the Appendix, the reform is likely to have opposite effects in rich and poor municipalities. In rich municipalities, where $E(z^a) > E(z^p)$, the reform certainly increases the expected welfare of the consumer when the incumbent is of type a and might increase her welfare even if the incumbent is of type p, provided that there are enough a types in the population of politicians (e.g. π is large enough). In poor municipalities, where $E(z^a) < E(z^p)$, the opposite occurs. If the future matters enough (λ is large), the expected welfare of the consumer certainly falls if the incumbent is of the p-type and there are few a types in the population of politicians (e.g. for $\pi \to 0$), and it might also fall, under the same condition, if the incumbent is of a a-type (as a p-type would still more likely to be re-elected).

Thus, differently from Brollo et al. (2010), in our model reducing transfers do not necessarily increase the utility of consumers, even if municipalities are compensated so as to receive in expected terms the same resources after the reform. Quite intuitively, a revenue compensated decentralization reform in poor municipalities do not provide great advantages to the consumer: the $p-$types are still on average more efficient in these municipalities, they are still more likely to be re-elected, and so the only effect of the reform is to make these politicians less able to bring home additional resources, thus damaging the consumer (in the second period). However, as stressed above, these effects also depend on the share of the different types of politicians in the population. To this we now turn.

2.3 Endogenous candidacy

So far we took π as given; the share of potential politicians of both types is fixed and does not depend on financing rules. But this is clearly unrealistic; the change in financing rules also changes the expected rents for both types for becoming a politician, and therefore one would expect that this also affects both the size and the composition of the set of potential politicians. Endogenous candidacy could be modeled in a number of ways; for our aims here, the following simple approach suffices. Thus, suppose that at time 0, that is, before period 1 begins, a citizen of type j is considering whether entering in the political arena or not. Suppose that there are n^j such potential candidates of type j, where n^j 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 he 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 $\left\{\overline{w}^j, 0\right\}$. Citizen j observes the realization of his wage w^j before deciding whether becoming a member of the set of potential candidates j; he also knows the expected 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 he just ignores the latter in taking his candidacy decision. The only cost for a citizen j of becoming a politician is that if he is elected, he has to rule, giving up his private wage for the time he serves as an elected official. The candidacy decision is taken at the end of time 0 and cannot be revised after. 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 enfolds as already described in the previous sections.

Under these assumptions, the choice of citizen j at the end of time 0 is quite simple; he will accept to join the set of politicians if the expected rents from doing so (in the case he is selected to become the incumbent politician at period 1) overcome the foregone wages; that is, provided than $E(R^j) \geq (1+\lambda)w^j$[6]. The ex-ante probability (computed at the beginning of time 0, before the realization of w^j) that a citizen of type j joins the political market is then $\frac{1}{\overline{w}^j}\frac{R^j}{(1+\lambda)}$, and as all js face the same distribution, the expected number of individuals of type j (equal to the realized number for large n^j) who join the political market is then $J = \frac{1}{\overline{w}^j}\frac{E(R^j)}{(1+\lambda)}n^j$, where $J = A, P$. It immediately follows that $\pi = \frac{A}{A+P} = \frac{\gamma R^a}{R^a+R^p}$, where $\gamma = \frac{n^a}{\overline{w}^a}\frac{\overline{w}^p+\overline{w}^a}{n^a+n^p}$.

In general, one would expect both $\overline{w}^j$ and n^j to differ among types, although it is hard to say how. It is likely that the market opportunities for the a-types are positively correlated with the tax base of the municipality, that it could be captured here with a higher $\overline{w}^a$ in municipalities with a higher α. But possibly, the same is true even for the p-type, although it is more difficult to make plausible conjectures in this case. Lacking further information and for the sake of simplicity, in the following we just impose $\gamma = 1$.

Notice from (**??**), 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 depend on the probability of meeting different types of opponents.

[6] Strictly speaking, this formulation is not correct, as it implicitely assumes that the consumer does not earn a wage in the second period even if he is not re-elected. Furthermore in his computation, the citizen should also consider the possibility that if he is not chosen as an incumbent in the first period he might be still chosen as an opponent in the second, then winning in this period with a probability that depends on the type of the incumbent, and thus gaining in this period a net benefit equal to $\overline{R} - w^j$. For simplicity, we ignore these further complications here, as they would greatly complicate the derivation of the formulas below. But note that considering them would not change qualitatively the results to follow, that is that richer municipalities should have a larger number of a-type incumbents. Details are available from the authors on request.

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 equilbrium π^*, these two forces need to balance. As the Appendix shows, solving the resulting system of simultaneous equations, this equilibrium share can be computed as:

$$\pi^* = \frac{1}{2} + \frac{(E(z^a) - E(z^p))(1 + \frac{\lambda \overline{R} \psi}{\tau})}{c(A+P)} \tag{8}$$

where $c = \frac{2\overline{w}(1+\lambda)}{n} > 0$. Thus, the denominator of the second term on the RHS of (8) is just proportional to the total number of potential politicians of both types and it is strictly positive. Equation (8) allows us to get an important conclusion:

Corollary 4 *$\alpha \geq (<) \alpha^*$ 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 type a. Using (8), we can also investigate the effect of the Italian reform on π^*.

Proposition 5 *$\partial \pi^* / \partial t > 0$ for $\pi^* \geq \frac{1}{2}$. $\partial \pi^* / \partial t > 0$ for π^* smaller but close to $\frac{1}{2}$. For $\pi^* \to 0$, the sign of $\partial \pi^* / \partial t$ is uncertain and might become negative.*

Thus, our model implies that the reform would have the effect of increasing even further the divergence between municipalities. After the reform richer municipalities should have even more politicians of the a-type, 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 (8), 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 8), 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 with π^* not too far from $\frac{1}{2}$, the first effect still dominate the second, so leading to an increase in π^* albeit at a reduced rate. Finally, for very 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. In rich communities, the reform increases directly the expected welfare of voters and might increase even further because it attracts

more efficient politicians in the political market. In poor communities, the reform reduces the efficiency level of the p type politicians and might also reduce the share of a type politicians, so leading to lower level of expected consumer's welfare.

Summing up, our model 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 municipalties. In poorer municipalities, the increase of the share of this type of politician 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. This occurs even with exogenous candidacy but the effect is emphasized by endogenous candidacy as poorer municipalities attract fewer politicians with high administrative skills. Bearing these predictions in mind, let us then examine our data.

3 Setting the stage: the Italian Municipalities and the reforms of the Nineties

Italian Municipalities are the layer of government closest to citizens in the Italian administration system, below Provinces and Regions, the other two decentralized levels of government. They are in charge of a number of services, ranging from general administrative services (like the management of the Registry Office) to environmental services (like garbage collection and disposal). Differences among the more than 8,000 Municipalities are wide, with respect to size, average income, population density and composition.

Two fundamental reforms changed the working of the Italian Municipalities during the Nineties: the reform of the electoral rules introduced the direct election of the Mayor, changing her role; the introduction of the Imposta Comunale sugli Immobili (ICI), the municipal property tax, heavily contributed to increase the level of fiscal autonomy granted to municipal governments.

3.1 The reform of the municipal electoral law

Before 1993, the Italian municipal elections were held under a pure parliamentary proportional system. Citizens voted for parties' representatives to elect the City Council (Consiglio Comunale). Elected representatives had then to form coalitions, and retained the power to choose the Mayor (Sindaco) and the members of the Executive Office (Assessori). This pure parliamentary proportional system was characterized by political instability and lack of both transparency and accountability, for two main reasons: 1) the City Council often ruled by changing majority coalitions, which implied also a replacement of the Mayor and of the members of the executive office during a legislature; 2) parties were not stimulated to form coalitions before the elections, as these were normally decided (and very often reshuffled) after the electoral round; hence, voters were

not able to choose a clear majority coalition and a specific, although incomplete, political program.

The reform of municipal elections, implemented by the law 81/1993, replaced the pure parliamentary proportional system with a presidential plurality rule system, which allowed citizens to vote directly for the Mayor since then. Within this general framework, the law established two different set of rules for two groups of Municipalities: the small ones, with less than 15,000 inhabitants, and all the others. As for the first group, the law provides a single ballot plurality rule: the mayoral candidate that gets more votes is directly elected. Each candidate can be supported by one coalition only, often formed by merging parties that usually recover their identities after the elections. Voters cast one single vote for a mayoral candidate and her supporting list. Two thirds of the seats in the City Council are assigned to the coalition supporting the winner mayoral candidate, while the remaining one third is proportionally distributed among the other coalitions, provided they gained at least 4% of votes. Given that each voter can assign only one preference, the candidates as town councillors that get more preferences are elected to the Council.

As for the second group, the Municipalities with more than 15,000 inhabitants (all those considered in our empirical analysis), the new rules establish a runoff (or dual ballot) system. Electors are allowed to vote for the Mayor and for the City Council, expressing only one preference for each choice. They can however vote disjointly: i.e., they can vote for Mayor A and, at the same time, for a candidate as town councillor of a list supporting Mayor B. Each candidate for mayor can be supported by more than one list. The mayoral candidate that gets more than 50% of the votes at the first round is immediately elected, with a distribution of seats in the City Council that provides a premium for the winning coalition. In particular, 60% of the seats is ensured to the lists connected to the winning Mayor, while the remaining 40% are proportionally allocated among the other lists, provided they gained at least 3% of votes. This majority premium is not assigned either when the lists supporting the winner do not obtain more than 40% of votes, or when the coalition supporting the looser obtains more than 50% of votes. In both these situations, seats are distributed proportionally. If no candidates get more than 50% of preferences in the first round, the two ones that obtain the relative majority of votes are admitted to the second ballot, in which a vote can be expressed for the Mayor only, and not for the Council lists. In this second ballot, the candidate that obtains the majority of votes is elected. Notice that, between the two rounds, the law allows parties to form new coalitions, with the lists supporting excluded candidates in the first round that can decide to endorse one of the remaining candidates. The majority premium (60% of the seats) is guaranteed also in this case to the parties and the lists that support the winning candidate, provided that the coalition of the loser did not obtain more than 50% of the votes in the first round; a situation that, again, requires a proportional allocation of the seats.

The reform also introduced a ceiling of two consecutive terms for the Mayor, after which one needs to wait at least a legislature to be allowed to run again

as mayoral candidate. The duration of the legislature has been changed several times during the years: it was 5 years before 1993, it has been reduced to 4 years in 1993, to be brought back to 5 years in 2000.

The new electoral system introduced in 1993 has proved to be better than the old one, in terms of more stable and more compact majorities and, consequently, in terms of governability. Furthermore, the new system clearly accentuated the importance of the Mayor. First, the Mayor retains the power to appoint the executive officers and the city managers. Second, differently from the past, the Mayor cannot be removed without going through a new round of elections. Indeed, with the new presidential system, if the majority in the Town Council decides to remove his support to the elected Mayor, new elections need to be held, and the legislature ends. This increased importance of the Mayor emphasises personal characteristics of candidates, with likely effects in terms of both parties selection and potential candidates decisions to run. As for parties, the reform should have provided them incentives to choose different candidates, more competent and more independent, perhaps more politically connected with national politicians (e.g., Pasquino, 2006). As for potential candidates, the reform should have offered more skilled candidates the opportunity to enter into the local political arena. Of course, these effects should have been the same in all Italian cities. The reform of funding introduced however different incentives for different types of municipalities.

3.2 The introduction of a municipal property tax

The second important reform for the working of Italian Municipalities implemented during the Nineties is the introduction of the Imposta Comunale sugli Immobili (ICI), a municipal property tax on the value of buildings and building lands, introduced by the Legislative Decree 504/1992. While the tax base is uniformly defined in all the local jurisdictions (considering cadastral rents for buildings and current market values for building lands), Municipalities retain a certain degree of autonomy in (a) setting the 'ordinary' tax rate and (b) applying specific rates and exemptions for buildings used as main dwellings. As for the 'ordinary' (or 'business') tax rate (to be applied to building lands, agricultural lands, holiday houses, offices, shops and all buildings not used as dwellings), Municipalities can choose inside a range between 0.4 and 0.7 percent.

This council property tax represented a sharp change in the level of fiscal autonomy granted to Municipalities, that were previously largely dependent from transfers by the Central government. In fact, transfers to Municipalities from 1993 onwards have been cut for an amount equal to estimated revenues from ICI setting the tax rate at the minimum 0.4 percent level.

According to available statistics, in the last years ICI accounted, on average, for more than 50% of total tax revenues of Municipalities, and more than 25% of their total spending (e.g., Padovano, 2008). However, these averages hide substantial differences between rich Municipalities (especially in central and northern Italy) and the less wealthy ones (especially in the South). After the introduction of ICI, rich Northern Municipalities covered almost 70% of

their total expenditures with their own revenues (which includes ICI, but also a surcharge on the Personal Income Tax), while this percentage in the South has never exceeded 30%, with a national average of more than 50% (e.g., Bordignon, 2000). Hence, the increase in fiscal autonomy was effective only for rich Municipalities in the Centre-North, while vertical imbalance was only slightly modified for Southern Municipalities, that still continued to be largely dependent from Central government transfers.

As the reform of the electoral rule, also the introduction of ICI contributed to change the Mayor's role, increasing her autonomy and her powers. Furthermore, since ICI is a tax clearly managed at the local level, also accountability should have improved: with respect to a situation in which funding was (almost) completely in the hands of the central government and citizens were not able to recognize whom to blame for the taxes they had to pay, the choice of the ICI tax rate makes the actions of the Mayors more transparent to electors, making easier for them to evaluate her performance. The increased autonomy and transparency of policy choices, should have made the task of Mayor more complex, but (perhaps) more attractive for skilled people, with high outside options and high opportunity costs, providing them incentives to enter into the local political arena, and to parties to select better candidates. But differently from the reform of the electoral rule, the introduction of ICI should have generated the right incentives only for wealthy municipalities.

4 The empirical analysis

4.1 The identification strategy

Our working hypothesis is that the reform of the electoral rule should have affected the 'quality' of local politicians in the same way in all the Municipalities across the country, while the introduction of ICI should have influenced the characteristics of the Mayors differently in richer and poorer Municipalities. To test this hypothesis we then consider a Difference-in-Differences approach, assuming that the 'treatment' is represented by the reduction in the degree of Vertical Fiscal Imbalance following the introduction of ICI, and the 'treated' are the richer Municipalities according to our theoretical interpretation above. Let's define $Y^j,\ j = a, p, w$, the indicators for $a-$type politicians, $p-$type politicians and consumers' welfare, respectively. The effects we are looking at can be defined as follows:

$$DiD = \left[\left(\overline{Y}_{after\_decentralization}|Rich\right) - \left(\overline{Y}_{before\_decentralization}|Rich\right)\right] + \\ - \left[\left(\overline{Y}_{after\_decentralization}|Poor\right) - \left(\overline{Y}_{before\_decentralization}|Poor\right)\right] \quad (9)$$

where *Rich* and *Poor* is an indicator - based, for instance, on municipal per capita GDP - that allows us to classify richer and poorer communities. Considering the theoretical model, we expect the share of $a-$type politicians to have increased more in richer Municipalities and $p-$type to have increased in

poorer ones after the introduction of the Municipal property tax. As for citizens' welfare, we expect it to increase more in richer municipalities. Hence, we expect DiD to be positive for Y^a and Y^w, while negative for Y^p. Identification of these effects is based on the large variation in income distribution across Italian municipalities.

Of course, simple differences may be due to factors other than improvements in fiscal decentralization. To estimate these differences between rich and poor municipalities taking into account potential confounding factors we recur to the following econometric model:

$$Y_{it}^j = \alpha + \beta_1 GDPpc_i + \beta_2 DECENTR_t + \beta_3 GDPpc_i \times DECENTR_t + \gamma \mathbf{X}_{it} + \varepsilon_{it} \tag{10}$$

where $GDPpc$ is the municipal per capita GDP, our variable to identify the 'treated' (see below); $DECENTR$ is a dummy variable to identify the 'treatment', and $\mathbf{X}$ a vector of political and economic variables to characterize each municipality. Our working hypotheses are tested by looking at the sign and the statistical significance of β_3.

4.2 Data and variables definition

The original database we use for this study contains detailed information on the Mayors of the 89 Chief Provincial Towns of the 15 Ordinary Statute Regions elected between 1988 and 1997, which are the ten years around the introduction of ICI. We also extend the time-span to the period between 1995 and 2002, to include the eight years around the introduction of the Personal Income Tax surcharge ('addizionale IRPEF', see below)[7]. The database has been built mainly starting from the archive provided by the Italian Ministry of Domestic Affairs. Despite its usefulness, information in this archive are based on self-declaration made by the same politicians. Hence, original information have been carefully revised using other sources, such as newspaper's online archives and internet sites. For instance, we identified the exact party of every Mayor, as often this information is replaced by other labels (e.g., civic list) that did not allow us to identify their correct political affiliation. Main information contained in the archive on the individual characteristics of the Mayors include sex, age, date and place of birth, party affiliation, the level of education, and the job before entering into politics. We exploit these data to define some of our indicators of the quality of local politicians.

The dependent variables. We follow Nannicini and Gagliarducci (2011) to build our indicators for the type of politicians Y^a and Y^p. In particular, the indicator for $a-$type politicians is a dummy taking value one when the elected mayor comes from a high skilled job. These jobs include white collars, entrepreneurs and directors; all past professions have been identified according

[7] Notice that we consider here eight years instead of ten, because the PIT surcharge was blocked after 2002.

to the classification provided by the Italian National Institute of Statistics (ISTAT). Since according to our theoretical model, the $a-$types are those endowed with "administrative" skills, we also experimented with alternative definitions of Y^a, considering more stringent definitions of jobs. The indicator for $p-$type politicians is built upon information on the political career of each Mayor. From the archive of the Ministry of Domestic Affairs, we derived information on the legislatures covered by each Mayor, considering the date of election, the expiring date, and the way the mandate ended (natural deadline, no-confidence vote, resignation). We added information on all the political offices covered by every first citizen in the sample during her life. The sources of these information are mainly four: the online registry office of the Italian Ministry of Domestic Affairs provided information about the political offices covered at the local level by every politician in Italy; the online historical archive of both the Italian Chamber of Deputies and the Senate of the Italian Republic (the two Chambers of the Italian Parliament) reports past career of all deputies and senators elected from the first legislature; the online archive of the European Parliament reports information of elected politicians. From these data, we first compute the years of 'political experience' before being elected as a Mayor, considering a political office in Italian or European institutions (local councils, local governments, national government, Italian Parliament or 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). This variable measures then the percentage of working years a person has dedicated to politics. Finally, the indicator for consumers' welfare Y^w is defined as the percentage of separate waste collection, an indicator provided by Legambiente, an Italian independent environmental organization.

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.

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 include two dummy variables, LN and FI, which takes value one when the mayor belong to the Lega Nord and to Forza Italia, respectively. Both the Lega Nord and Forza Italia are two parties originated from the abrupt change which characterized the Italian party system at the beginning of the '90s, caused by the scandal of Mani Pulite (the prosecution of a number of politicians charged of corruption in public affairs). We also include, but only in the model considering citizens' welfare, a dummy taking value one if the mayor is subject to a term limit according to Italian legislation (TL). Albeit rough, this is a control allowing us to separate the 'selection' effect from the 'opportunistic behaviour' effect. As for the political characteristics of the municipality, we first consider the dummy variable $ALLINEAMENTO$, 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). We then consider an index of political competition, defining the variable $COMPETITION$ as the Herfindahl index on the shares of Centre-Left and Centre-Right coalitions. As suggested by, e.g., Nannicini and

Galasso (2011), competition could have an important impact on the selection of politicians.

Socio-economic 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 (see, e.g., Nannicini and Gagliarducci, 2011). To take into account the characteristics of the potential pool of candidates which 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 ($ANZIANI\_PERC$). We also consider $MOVIPROV$, the number of enterprises relative to total population, as a measure of the oppotunity cost of entering into politics. Finally, we control for macro-area dummies as a control for structural differences across different municipalities, such as for instance the social capital.

4.3 Results

Before moving to a more formal econometric analysis, let us first consider simple difference-in-differences in the three output indicators. To this end we define poor municipalities those with per capita GDP below the median, and as rich municipalities those with per capita income above the median. Tables 1.a to 1.c show evidence supporting our working hypothesis. As for the indicator for $a-$type politicians (Table 1.a), the descriptive evidence suggests a difference-in-differences of 33.8%: after introducing ICI, the increase in the share of $a-$type politicians was much more marked in rich municipalities than in poorer ones (40.9% vs 7.1%). The same is true for the indicator of consumers' welfare (Table 1.c): the increase for richer municipalities was 2.4% vs 0.9% for poorer ones, with a difference-in-differences of 1.5%. On the contrary, Y^p reduced more in richer communities than in poorer ones (-7.7 vs -3.3 in the past political experience ratio), with a difference in differences of -4.4 (Table 1.b).

Of course, these simple differences might be due to factors completely different from the introduction of ICI, like - for instance - a higher endowment of social capital for richer municipalities with respect to poorer ones, or a prevalence of candidates from parties borne from the corruption scandal Mani Pulite (like Lega Nord and Forza Italia) again in richer contexts. To control for these and other potential confounding factors, we estimate Equation (10) with OLS, using Panel Corrected Standard Errors to control for heteroskedasticity and contemporaneous cross-section correlation among municipalities (see, e.g., Hoechle, 2007). Estimates of our empirical model in Tables 2.a to 2.c largely confirms the simple descriptive evidence provided above. Estimated $\widehat{\beta}_3$ take up the expected sign (positive for Y^a and Y^w, negative for Y^p), and are statistically significant at the usual confidence levels. Interesting patterns emerge also for other controls. For instance, the two dummies for Lega Nord and Forza Italia are both negative and statistically significant for the indicator of $p-$type politicians, suggesting that these new parties do indeed select less politically experienced

candidates to run for mayor (Table 2.b). However, only Forza Italia selected more $a-$type politicians: the estimated coefficient is positive and statistically significant, while the one for Lega Nord is not (Table 2.a). On the contrary, both dummies are statistically insignificant in the consumers' welfare equation (Table 2.c).

4.4 Robustness checks

We provide a number of robustness checks for our results.

- A first problem is that the decentralization process may interact also with other political variables, like the degree of political competition and the alignment between local and central government. Previous results are however robust to the inclusion in the model of additional interaction terms ($ALLINEAMENTO \times GDP$; $ALLINEAMENTO \times DECENTR$; $COMPETITION \times GDP$; $COMPETITION \times DECENTR$).

- A second, more important, problem is that decentralization was implemented together with changes in the electoral rules. We are unable to check for the role played by these new rules because the dummy $DECENTR$ takes value one the first time Municipalities voted after the introduction of ICI in 1993 *with the new rules*. We then consider a different decentralization experiment: the introduction of the Personal Income Tax surcharge in 1999. Considering a panel on the 1995-2002 period, we then estimate our previous Equation (10), where the dummy $DECENTR$ takes value one from 1999 onwards. Again, our estimates largely confirm previous result but for the indicator of $a-$type politicians as defined above, despite the fact that the reduction in the Vertical Fiscal Imbalance following the introduction of the PIT surcharge was lower than the reduction experimented with the introduction of ICI. For Y^a we experimented with different definitions. Results suggest that "administratively-skilled" individuals might be different from "high-skilled" ones.

- A third final control is related to resources for political campaigns. After the introduction of the new electoral rules, candidates spend more money for their electoral campaigns: richer candidates have more chances to win where income per capita is higher. Despite being required by the law, data on campaigning costs are unavailable. We argue that this problem is more severe in regional chief towns. We then re-estimate our models by dropping regional chief towns: results are largely confirmed also in this case.

5 Concluding remarks

In this paper we suggest that fiscal federalism works when the degree of Vertical Fiscal Imbalance is low. This is because the "quality" of decentralization goes

hand-in-hand with the "quality" of politicians: where the degree of Vertical Fiscal Imbalance is low, more "administratively-skilled" politicians are in office. We first provide a theoretical framework to show how this mechanism works, and an empirical analysis that largely confirms the theoretical predictions based on the largest reduction in the degree of Vertical Fiscal Imbalance for Italian municipalities, namely the introduction of the Municipal property tax in 1993. We also show that these empirical results are robust to other relevant stories, such us the possible interactions of our "treatment" variables with other political variables, the implementation of the new electoral rules in the same year of the Municipal property tax, and the different resources for political campaign for different candidates.

This paper provides a general insight on why fiscal federalism does not work when Vertical Fiscal Imbalance is high, and suggests that - in Italy - the problem for most public administrations in the South of the country is not a lack of social capital, but simply the rational behavior of voters. An important policy implication is that to make fiscal federalism work when Vertical Fiscal Imbalance is high, we need extra control on politicians more than voting. This could be an avenue to support a move toward some forms of "asymmetric federalism".

References

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

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

[3] Besley T., Coate S. (1997). "An economic model of representative democracy." Quarterly Journal of Economics, 112: 85-114.

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

[5] Bordignon M. (2000). "Problems of soft-budget constraints in intergovermental relationship: the case of Italy." No 3099, RES Working Papers from Inter-American Development Bank, Research Department.

[6] Bordignon M., Piazza S. (2010). "Who do you Blame in Local Finance? An Analysis of Municipal Financing in Italy". CESifo Working Paper Series, n. 3100.

[7] Bordignon M., Turati G. (2009). "Bailing out expectations and public health expenditure". Journal of Health Economics, 28(2): 305-321.

[8] Brollo F., Nannicini T., Perotti R., Tabellini G. (2010). "The Political Resource Curse". NBER Working Papers n. 15705.

[9] Carrillo J. D., Mariotti T. (2001). “Electoral competition and politician turnover” European Economic Review, 45: 1-25.

[10] Dahlberg M., Mörk E., Rattsø J, Ågren H. (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.

[11] Eyraud L., Lusinyan L. (2011). “Decentralizing Spending More than Revenue: Does It Hurt Fiscal Performance?”, IMF Working Paper, n. 226.

[12] Ferraz C., Finan F. (2009). “Motivating Politicians: the impacts of monetary incentives on Quality and Performance.” NBER Working Paper No. 14906.

[13] Guiso L., Sapienza P., Zingales L. (2010). "Civic Capital as the Missing Link". NBER Working Papers n. 15845.

[14] Hoechle D. (2007). “Robust standard errors for panel regressions with cross-sectional dependence” Stata Journal, 7(3): 281-312.

[15] Lockwood B. (2006). “Fiscal Decentralization: A Political Economy Perspective”, in E. Ahmad and G. Brosio (eds.), The Handbook of Fiscal Federalism, Edward Elgar.

[16] 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.

[17] Merlo A., Galasso V., Landi M., Mattozzi A. (2009). “The Labor Market of Italian Politicians.” PIER Working Paper No. 09-024.

[18] Merlo A., Mattozzi A. (2007). “The Transparency of Politics and the Quality of Politicians.” American Economic Review, 97(2): 311-315.

[19] Merlo A., Mattozzi A. (2010). “Mediocracy.” PIER Working Paper No. 10-017.

[20] Nannicini T., Gagliarducci S. (2011). “Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection”. Journal of the European Economic Association, forthcoming.

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

[22] Osborne M. J., Slivinski A. (1996). “A model of political competition with citizen candidates.” Quarterly Journal of Economics, 111(1): 65–96.

[23] Padovano F. (2008). “Setting House Taxes by Italian Municipalities: What the Data Say”, in F. Padovano, R. Ricciuti (eds.) Italian Institutional Reforms: A Public Choice Perspective. New York, Springer: pp: 89-115.

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

[25] Poutvaara P., Kotakorpi K. (2009). "Pay for Politicians and Candidate Selection: An Empirical Analysis." IZA Discussion Paper No. 4235.

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

[27] Rodden J. (2005). "And the Last Shall be First: Federalism and Soft Budget Constraints in Germany", MIT, mimeo.

[28] Veronese B. (2004). "Representation and institutional change". Sticerd Working Paper, London School of Economics.

6 Appendix

6.1 Proof of Proposition 1

As stated in the text, 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 $E(R^{j^*}) \geq \overline{R}$ for both $j = a, p$. Let then $E(z^k) = \min(\ E(z^a); E(z^p))$; the candidate more willing to deviate is then type k. For this candidate not to deviate, it must then hold:

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

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

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

Notice from (7), that (7) 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:

$$\widetilde{U} = E(z^k) - \overline{R} + \lambda \left[E(z^o) - \overline{R}\right] \tag{12}$$

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

$$U^* = \frac{\delta \overline{R} \psi}{\tau} E(z^o) + \lambda [\frac{s}{2}(E(z^k) + E(z^k | z^k \geq E(z^k)) + \tag{13}$$
$$+(1-s)((1-h^k)E(z^{-k}) + h^k E(z^k | z^k \geq E(z^{-k}))] - \lambda \overline{R}$$

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

$$\overline{R} \geq \frac{E(z^k)}{(\frac{\lambda \psi}{\tau} E(z^o) + 1)} = R'' \tag{14}$$

Combining (11) and (14), we see that $R' > R''$ if $m = \frac{1}{2s[E(z^{-k}) - E(z^k)]} > \frac{\psi}{\tau}$, which is already implied by A.2. Hence, $R' > R''$. As stated in Proposition 1, for $R' > \overline{R} > R''$ we then get an equilibrium. QED.

6.2 Proof of Lemma 1

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

$$U^a = \frac{\delta \overline{R} \psi}{\tau} E(z^o) - \lambda \overline{R} + \lambda (\frac{1}{2}(E(z^o) + \pi E(z^a | z^a \geq E(z^a)) + (1-\pi) E(z^a | z^a \geq E(z^p))) + \tag{15}$$
$$+\lambda (1-\pi) \frac{\psi}{\tau} \left[E(z^a) - E(z^p)\right] \left(E(z^a | z^a \geq E(z^p)) - E(z^p)\right);$$

$$U^p = \frac{\delta \overline{R} \psi}{\tau} E(z^o) - \lambda \overline{R} + \lambda (\frac{1}{2}(E(z^o) + (1-\pi) E(z^p | z^p \geq E(z^p)) + \pi E(z^p | z^p \geq E(z^a))) + \tag{16}$$
$$+\lambda \pi \frac{\psi}{\tau} \left[E(z^p) - E(z^a)\right] \left(E(z^p | z^p \geq E(z^a)) - E(z^a)\right);$$

for the a-type and the p-type, respectively.

Invoking (3), and solving the integrals we obtain:

$$E(z^a | z^a \geq E(z^a)) = \frac{E(z^a)}{2} + T = x$$

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

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

$$E(z^p | z^p \geq E(z^p)) = \frac{E(z^p)}{2} + T = x - \frac{h}{2}$$

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

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

where $T = \frac{3\tau^2+a^2t^2}{24\tau\psi}$, $h = E(z^a) - E(z^p)$, $y = (E(z^a) + E(z^p))$. Notice that $T > \max(\frac{E(z^a)}{2}; \frac{E(z^p)}{2})$. This also implies $2T > y$.

Substituting in (15) and (16) and simplifying:

$$\begin{aligned} U^a &= \frac{\delta\overline{R}\psi}{\tau}E(z^o) - \lambda\overline{R} + \lambda(\frac{1}{2}(E(z^o) + \frac{E(z^a)}{2} + T) + \\ &\quad +(1-\pi)\frac{\psi}{\tau}h(T + \frac{h - E(z^p)}{2} + \frac{1}{2}y(\frac{1}{2} + \frac{\psi}{\tau}h)) \end{aligned}$$

$$\begin{aligned} U^p &= \frac{\delta\overline{R}\psi}{\tau}E(z^o) - \lambda\overline{R} + \lambda(\frac{1}{2}(E(z^o) + \frac{E(z^p)}{2} + T) + \\ &\quad +\pi\frac{\psi}{\tau}h(-T + \frac{h + E(z^a)}{2} + \frac{1}{2}y(\frac{1}{2} + \frac{\psi}{\tau}h)) \end{aligned}$$

It follows:

$$U^a - U^p = \lambda h\left[\frac{1}{2} + \frac{\psi}{\tau}(T + \frac{h - E(z^p)}{2} + \frac{1}{2}y(\frac{1}{2} + \frac{\psi}{\tau}h) - 2\pi(2T - y))\right] \quad (17)$$

Inspection of the term in square brackets on the RHS of (17) shows that this term is always positive and decreasing in π. It follows: i) sign$(U^a - U^p)$=sign$(E(z^a) - E(z^p))$; ii) $(U^a - U^p)$ if positive (negative), decreases (increases) in π. QED.

6.3 Proof of Proposition 2

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

- for $a-$type

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

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

We have:

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

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

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

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

which is positive if $\theta^a(1+\lambda\overline{R}\frac{\psi}{\tau^2}(1-2\pi)) > \lambda^a + \lambda\overline{R}\frac{\psi a}{\tau^2}2(1-\pi)\theta^p$.
This is sure if $\theta^a \geq \frac{2(1-\pi)}{(1-2\pi)}\theta^p$; the worst possible case is $\pi = 1$.

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

Consider now

$$\frac{d^2}{dtda}E(R^a) = (\theta^a-\delta^a)+\lambda\overline{R}\frac{\psi}{\tau^2}((1-2\pi)(\theta^a-\theta^p)-\theta^p)+2\lambda\overline{R}\frac{\psi at}{\tau^3}((1-2\pi)(\theta^a-\theta^p)-\theta^p)$$

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

which is positive if $\theta^a(1+\lambda\overline{R}\frac{\psi}{\tau^2}\frac{(1+at)}{(1-at)}((1-2\pi)) > \delta^a + \lambda\overline{R}\frac{\psi a}{\tau^2}2(1-\pi)\theta^p$.
This is sure if $\theta^a \geq \frac{2(1-\pi)}{(1-2\pi)}\theta^p$; the worst possible case is $\pi = 1$.

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

- For $p-$type:

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

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

we have

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

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

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

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

and

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

Finally:

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

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

QED.

6.4 Proof of Proposition 3

Consider again, the welfare of the consumer under the two types:

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

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

where $T = \frac{3\tau^2+a^2t^2}{24\tau\psi}$, $h = E(z^a) - E(z^p)$, $y = (E(z^a) + E(z^p))$. Note that the terms in square brackets in the RHS of both equations are positive.

Differentiating the two expression above with respect to t produce effects that are generally uncertain. Notice however that $\frac{d}{dt}(\frac{\delta\overline{R}\psi}{\tau}E(z^o)) = \frac{\delta\overline{R}\psi}{\tau^2}\alpha(\pi\overline{\theta}^a + (1-\pi)\overline{\theta}^p) > 0$, so that the first period welfare effects are positive and larger in richer municipalities. Further note, $\partial T/\partial t = \frac{\alpha(-3\tau^2+\alpha^2t^2)}{24\psi\tau^2} < 0$, so that $\frac{d}{dt}(\frac{E(z^p)}{2} + T) < 0$ while $\frac{d}{dt}(\frac{E(z^a)}{2} + T)$ may be positive. This implies that the derivative wrt t of the third term is more positive or less negative for consumer's welfare under type a. Finally, $\frac{d}{dt}h > 0$, while $\frac{d}{dt}y$ is positive (negative) if $\theta^a - \delta^a > (<)\theta^p - \delta^p$. Putting together these observations, some results emerge for special case. Suppose for instance $\pi \to 1$. Then, $\frac{d}{dt}U^a > 0$, while $\frac{d}{dt}U^p$ is uncertain. Suppose $\pi \to 0$. Then $\frac{d}{dt}U^p < 0$ for λ large enough, while $\frac{d}{dt}U^a$ is uncertain. By continuity, these results also hold for π not too far from the extremes. For $1 > \pi > 0$ effects are uncertain. However, for $h > 0$ and $\frac{d}{dt}y > 0$, the derivative of the fourth term wrt t is positive in both cases, meaning that the welfare of the consumer increases as t increases. For $h < 0$ the effect is uncertain, and the derivative of the fourth term may become negative for $\frac{d}{dt}y < 0$. The above suggests the following conjecture. In rich municipalities, where $h > 0$ the reform certainly increases the welfare of the consumer when the incumbent is of type a and might increase her welfare even if the incumbent is of type p, provided that there are enough a types in the population of politicians (e.g. π is large enough). In poor municipalities, where $h < 0$, the reform might instead reduce the welfare of the consumer when the incumbent is of type p, and this is particularly the case when there are few a types in the population of politicians (e.g. π is small enough). QED

6.5 Derivation of π^*

Let:

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

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

where $\pi = \frac{E(R^{a^*})}{E(R^{p^*})+E(R^{a^*})}$.
Rewriting and simplifying:

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

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

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

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

Also note:

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

Substituting:

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

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

Let $A + B = x$; $\frac{1}{2}\lambda \overline{R} + (a+p)(1-k) = b; 2(a-p)^2 k(1+k) = c$, leading to $x^2 - bx + c = 0$. Solving, the two roots are $\frac{1}{2}b + \frac{1}{2}\sqrt{b^2 - 4c}; \frac{1}{2}b - \frac{1}{2}\sqrt{b^2 - 4c}$. The equation admits real solutions if $b^2 \geq 4c$; this is certainly the case for $|a-p| \leq \frac{b}{2\sqrt{2(k+k^2)}} = Q$. Note that Q is decreasing in k and that for $k \to 1$, $Q \to \frac{\tau}{8\psi}$. 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 $a \to p$. Thus, the only economic significant solution is represented by the positive root. Under the positive root, $0 \leq \pi \leq 1$ iff $|a-p| \leq \frac{A+B}{(1+k)} = T$. Furthermore, under the positive root, $A+B$ is increasing in $(a+p)$ and decreasing in $|a-p|$ and in k. The highest value of $A+B$ is b for $a = p$. It follows that $T \leq Q$ for k sufficiently smaller of 1. Finally, note that $A + P = \frac{n(b+\sqrt{b^2-4c))}}{2w(1+\lambda)}$. Substituting this gives equation (8) into the main text. QED.

6.6 Proof of Proposition 5

Differentiating the numerator and the denominator of (12) for t and imposing $\tau = 1 - \alpha t$, we get:

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

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

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

for $E(z^a) \geq (<)E(z^p)$

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

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

It follows:

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

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

Table 1.a

MAYORS FROM HIGH SKILLED OCCUPATION (%)					
		Decentralization			
Rich					
		Before	After		Difference
No		42.9	50		7.1
Yes		26.5	67.4		40.9
Difference		-16.4	17.4		33.8

Table 1.b

PAST POLITICAL EXPERIENCE RATIO (%)					
		Decentralization			
Rich					
		Before	After		Difference
No		23.6	20.3		-3.3
Yes		28.2	20.5		-7.7
Difference		4.6	0.2		-4.4

Table 1.c

SEPARATE WASTE COLLECTION (%)					
		Decentralization			
Rich					
		Before	After		Difference
No		1.9	2.8		0.9
Yes		7.6	10.0		2.4
Difference		5.7	7.2		1.5

Table 2.a – a-type politicians

Linear regression, correlated panels corrected standard errors (PCSEs)

Group variable:	comune	Number of obs	=	760
Time variable:	anno	Number of groups	=	76
Panels:	correlated (balanced)	Obs per group: min	=	10
Autocorrelation:	no autocorrelation	avg	=	10
		max	=	10
Estimated covariances	= 2926	R-squared	=	0.1585
Estimated autocorrelations	= 0	Wald chi2(15)	=	445.22
Estimated coefficients	= 16	Prob > chi2	=	0.0000

quality	Panel-corrected Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
decen	-.2082021	.1166551	-1.78	0.074	-.4368419	.0204378
pil2010_m	-.0079822	.0591524	-0.13	0.893	-.1239188	.1079544
decen_pil	**.2353158**	**.051732**	**4.55**	**0.000**	**.1339229**	**.3367088**
sf	-.1926351	.078078	-2.47	0.014	-.3456652	-.0396051
age	.006838	.0018561	3.68	0.000	.0032	.0104759
pop	-4.94e-08	8.10e-08	-0.61	0.542	-2.08e-07	1.09e-07
anziani_perc	-.0278279	.0057111	-4.87	0.000	-.0390215	-.0166343
moviprov	-.0321033	.018965	-1.69	0.091	-.0692741	.0050675
nordovest	.1719912	.0691336	2.49	0.013	.0364919	.3074905
nordest	.1124362	.0393801	2.86	0.004	.0352527	.1896197
centro	.1416686	.0458277	3.09	0.002	.0518479	.2314892
allineamento	.0486075	.041788	1.16	0.245	-.0332954	.1305105
competition	-.0000633	.0000522	-1.21	0.225	-.0001655	.000039
LN	-.0479282	.1104729	-0.43	0.664	-.2644512	.1685948
FI	.4642113	.0657636	7.06	0.000	.335317	.5931057
_cons	.9079487	.3048189	2.98	0.003	.3105147	1.505383

Table 2.b – p-type politicians

```
Linear regression, correlated panels corrected standard errors (PCSEs)

Group variable:   comune                        Number of obs      =       885
Time variable:    anno                          Number of groups   =        89
Panels:           correlated (unbalanced)       Obs per group: min =         9
Autocorrelation:  no autocorrelation                           avg =   9.94382
Sigma computed by casewise selection                           max =        10
Estimated covariances      =      4005          R-squared          =    0.1426
Estimated autocorrelations =         0          Wald chi2(15)      =   5682.27
Estimated coefficients     =        16          Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |            Panel-corrected
perc_polit~a |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       decen |   .0719855     .04367     1.65   0.099    -.0136061    .1575771
   pil2010_m |   .0527605    .042327     1.25   0.213    -.0301989    .1357199
   decen_pil |  -.0453007   .0203245    -2.23   0.026     -.085136   -.0054653
          sf |  -.0590277    .034569    -1.71   0.088    -.1267817    .0087262
         age |   -.003659   .0007399    -4.95   0.000    -.0051092   -.0022089
         pop |   1.02e-08   2.07e-08     0.49   0.622    -3.04e-08    5.08e-08
anziani_perc |   .0013813   .0030728     0.45   0.653    -.0046413    .0074038
    moviprov |   -.000417   .0057373    -0.07   0.942     -.011662    .0108279
   nordovest |   .0362438   .0293577     1.23   0.217    -.0212963    .0937839
     nordest |   .0365301   .0407294     0.90   0.370     -.043298    .1163583
      centro |   .0425703   .0144073     2.95   0.003     .0143326     .070808
allineamento |   .0337329   .0157586     2.14   0.032     .0028467    .0646191
 competition |   .0000435   .0000403     1.08   0.281    -.0000355    .0001224
          LN |   -.183242   .0299812    -6.11   0.000     -.242004   -.1244799
          FI |  -.1546682   .0225154    -6.87   0.000    -.1987975   -.1105389
       _cons |   .0386316   .1736798     0.22   0.824    -.3017745    .3790377
------------------------------------------------------------------------------
```

Table 2.c – Consumers' welfare

Linear regression, correlated panels corrected standard errors (PCSEs)

Group variable:	comune	Number of obs	383
Time variable:	anno	Number of groups	83
Panels:	correlated (unbalanced)	Obs per group: min	2
Autocorrelation:	no autocorrelation	avg	4.614458
Sigma computed by	casewise selection	max	5
Estimated covariances	3486	R-squared	0.4852
Estimated autocorrelations	0	Wald chi2(5)	778.09
Estimated coefficients	17	Prob > chi2	0.0000

waste	Panel-corrected Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
decen	-3.227789	1.717433	-1.88	0.060	-6.593896	.1383178
pil2010_m	4.875703	1.059344	4.60	0.000	2.799427	6.951979
decen_pil	2.492204	1.195402	2.08	0.037	.1492594	4.835148
sf	-.3566804	1.126562	-0.32	0.752	-2.564701	1.85134
age	-.043933	.0109959	-4.00	0.000	-.0654846	-.0223814
pop	-3.10e-06	4.20e-07	-7.39	0.000	-3.93e-06	-2.28e-06
anziani_perc	-.2106057	.021042	-10.01	0.000	-.2518472	-.1693642
moviprov	.5790696	.2786273	2.08	0.038	.0329702	1.125169
nordovest	4.989052	.4078616	12.23	0.000	4.189658	5.788446
nordest	-.2880062	.7261374	-0.40	0.692	-1.711209	1.135197
centro	-.0410635	.6686958	-0.06	0.951	-1.351683	1.269556
allineamento	.9394668	.2829589	3.32	0.001	.3848776	1.494056
competition	.001931	.0008142	2.37	0.018	.0003352	.0035268
LN	.2616393	.6086883	0.43	0.667	-.9313679	1.454646
FI	-.2336937	.2226168	-1.05	0.294	-.6700147	.2026273
tl	.5762908	.6841166	0.84	0.400	-.764553	1.917135
_cons	-14.42852	4.580634	-3.15	0.002	-23.40639	-5.450639