Follow the herd. Spatial interactions in tax setting behaviour of Italian municipalities

Domenico Depalo* and Giovanna Messina †‡

August 18, 2011

Preliminary and incomplete

Abstract

This paper seeks to shed light on the determinants of local tax rates in Italy by focusing on the mutual interactions between local fiscal policies and the features of the electoral system, which is the key insight of the yardstick competition theory. In a framework of asymmetric information, voters use tax rates set by adjacent municipalities as a yardstick to evaluate the performance of their incumbents. Local administrators anticipate this behaviour and move as a herd, mimicking each other. Within this analytical framework we study the extent and the sign of spatial correlation among Italian municipal property tax rates in the years 1998-2006. We follow a spatial panel approach in order to control for individual time-invariant characteristics and gauge the extent of genuine tax mimicking behaviour. Our results are supportive of positive spatial correlation among property tax rates. Own tax rates appear to react more strongly to neighbours policy choices than to standard determinants of tax setting decisions. Our conviction that tax herding is driven by information spillovers rather than resource flows comes from the low mobility of the tax base and from a counterfactual evidence of tax rates increasing over time. Furthermore, such evidence calls for measures aimed at containing the risk of tax collusion that may arise from herding behaviour.

*Bank of Italy, Rome Branch, Regional Research Division

†Bank of Italy, Research Department

‡We would like to thank Roberto Basile, Valter Di Giacinto, Vanni Mengotto, Federico Revelli, Marzia Romanelli and Paolo Sestito for their useful comments. We thank seminar participants at the Brown Bag Lunch Meeting (Italian Ministry of Economic and Finance, Rome) and the V^{th} World Conference of the Spatial Econometric Association (Toulouse, France). All the codes are (minor) modifications from LeSage econometric toolbox and are available at the webpage: *http://sites.google.com/site/domdepalo/*. The views expressed in this paper are those of the author and do not imply any responsibility of the Bank of Italy. Corresponding author: giovanna.messina@bancaditalia.it

JEL classification: H71, D72.

Keywords: yardstick competition, electoral budget cycle, spatial econometrics .

1 Introduction

Claims for increasing fiscal powers of subnational governments have become a hot topic in Italy's everyday politics. A central tenet is that intergovernmental competition would spur efficiency and improve incentives for governments to act in their citizens' interest. But so far the debate is going without a supportive evidence on the determinants of local budgeting processes at the micro level, especially when taxing powers are concerned.

Our work seeks to shed light on the determinants of tax setting decisions concerning the property tax, which is the main own revenue of Italian municipalities, in a sample of more than 6.000 jurisdictions in the 1998-2006 period. Our research questions are: do governments engage in some form of strategic interaction when setting tax rates? Has this anything to do with politics? These questions are at the core of the yardstick competition (YC) theory. The YC theory was first introduced by Salmon (1987), who adapted to the field of public economics some analytical tools borrowed from the theory of regulations (Shleifer, 1985) and that of tournaments (Lazear and Rosen, 1981 and Nalebuff and Stiglitz, 1983). A more formal approach was subsequently developed by Besley and Case (1995a) in a study of governors' elections and national taxes in US states. The basic idea of YC is that citizens and politicians are linked through a typical principal-agent relationship. Since information on the cost of public services is asymmetric (only local officials are aware of it), voters are not able to discriminate if observed taxing decisions are necessary or are instead the result of rent extraction. Hence they use adjacent jurisdictions as a yardstick and defeat incumbents whose behaviour depart significantly from that of their neighbours.

YC models are empirically appealing since they produce a clear testable prediction, namely that tax rates are spatially correlated and that the source of tax mimicking is deeply rooted in the political contest. This issue has triggered a number of country-specific studies in the last decade. Some examples are the papers by Heyndels and Vuchelen (1998), Brett and Pinske (2000), Brueckner and Saavedra (2001) and Bordignon et al. (2003), which is also the only reference study for our country.

Results of spatial interactions among local fiscal policies have to be interpreted carefully. First, one has to be sure that such interactions capture effective strategic behaviour and not spatial dependence among the error terms (as in the case of omitted variables being themselves spatially correlated). Second, it has to be ascertained which are their underlying determinants, since both YC models and standard tax competition (TC) models predict a non zero slope in the reaction function of own tax rates to neighbours tax setting decisions.

To overcome the first drawback, we adopt a spatial panel approach (whereas most of empirical analyisis is focussed on the cross-section, with the exception of Revelli, 2001). This technique allows us to estimate a spatial coefficient of 0.10-0.12, which is most likely related to genuine mimicking behaviour and not to spatial error dependence due to mis-specification. The coefficient is lower than previous cross-section evidence on the Italian case, but yet it denotes a stronger reaction to neighbours policy choices than to standard determinants of tax setting decisions.

As for the nature of the driving process, we are confident that tax herding is of the yardstick type rather than arising from standard tax competition based on resource flows. This conviction comes from the intrinsic characteristics of the property tax base (which displays low mobility, especially in the Italian framework) and from a counterfactual evidence of increasing tax rates over time (whereas with TC one would expect decreasing or at least stable tax rates).

Such evidence casts some doubts about the welfare enhancing effect of tax decentralisation *via* improved competition among local governments. Indeed Besley and Smart (2002) have warned that information spillovers may perversely induce bad incumbents who fear to be removed from office to increase rent extraction up to the maximum. Hence it seems important that further steps towards tax decentralisation in Italy are complemented with a redesign of the structure of incentives faced by subnational governments. Measures such as more information on the features of the production function of local public goods as well as a proper design of electoral rules and institutions would contribute to map more readily citizens' preferences into local officials budgeting decisions.

The rest of the paper is organised as follows. Section2 describes the institutional framework and gives some stylised evidence about property tax rates; Section 3 presents an economic model; Section 4 describes our empirical strategy; Section 5 gives a short description of the data set. Section 6 describes the results. Section 7 argues on the sources of strategic interactions. Section 8 concludes.

2 Tax autonomy at municipal level: some stylised evidence

Italy is organized as a three-tier system of subnational governments, whose layers are respectively regions, provinces and municipalities. The latter are about 8.100 and are very small (on average each of them lay on area of 14 square miles). This makes Italian municipalities an ideal laboratory to test for spatial interactions and informational externalities, which intuitively are the more likely to arise the smaller is the scale of the units under observation.

Municipalities are multipurpose governments, their major functions being registry, waste disposal, urban planning, public illumination, road maintenance, local transports, social aid, childcare, primary schooling, and assistance to the elderly. Expenditures are financed by grants (roughly by 45 per cent), own taxes (30 per cent) and other own revenues (such as tariffs, fees and penalties).[1] Municipal taxes are manifold, but the major one is undoubtedly the local property tax (*Imposta comunale sugli immobili, ICI*), which accounts for more than half of total tax revenue (see Table 1); other important sources of revenues are the tax on solid waste management (which works fairly as a tariff rather than a tax, covering a given share of waste collection costs) and the surcharge on the national personal income tax.

Table 1: Composition of municipal tax revenues

Municipal Taxes	Share
Property tax	50.1%
Waste management	24.2%
Surcharge on personal income tax	14.3%
Surcharge on electricity consumption	3.6%
Advertising and occupation of public areas	3.2%
Other	4.6%

Source: *Relazione sul federalismo fiscale della Copaff* (30 June 2010, Appendix 2, p. 88.).

Our analysis concentrates on property tax rates, which are the substance of municipal fiscal powers. Such rates can be fixed in a range from 0.4 to 0.7 percentage points on the tax base and are usually differentiated between main residential properties (*ICI prima casa*), which also benefit of some allowances set at the local level, and business properties (*ICI ordinaria* including investment houses).[2] Our empirical analysis focuses on business properties rather than residential properties for several reasons. First, *ICI ordinaria* has a preminent role in financing municipalities, since it provided more than three quarters of total property tax revenue in the years from 1998 to 2006. Setting business property tax rates was probably the only expression of municipal fiscal powers in the period under observation since alternative forms of taxing powers, namely the ability to manoeuvre the surcharge on national personal income tax, were hampered by national legislation.[3] Finally, the analysis of business properties tax rates is more insightful in perspective, since by 2008 *ICI prima casa* has de facto been abolished and proposals for a future revision of municipal financing structure point to business properties as the principal source for local taxation.

[1] Overall municipalities account for almost 10 per cent of total general government expenditure net of interests

[2] The base of the property tax is the assessed rental value of immovable property as defined by central government for each cadastral category. Hence the tax base is determined uniformly across all jurisdictions

[3] The surcharge on national personal income tax was introduced in 1998 and municipalities were entitled to vary tax rates starting from 2001, but from September 2002 to the end of 2006 the ability to raise such tax rates was suspended by central government.

Descriptive evidence drawn from our data set shows that business property tax rates have raised from 0.543 percentage points in 1998 to 0.603 in 2006 (see Table 8). This trend seems rather uniform across country, despite the huge differences in economic conditions among Northern and Southern Italy. In particular, in the North-West area rates have steadily grown from 0.561 to 0.620, in North-East from 0.557 to 0.639, in Centre from 0.619 to 0.675, while in the South and in the Islands rates have raised from 0.561 to 0.643 and from 0.510 to 0.621 respectively. Apparently, there is not a clear financial explanation for such a trend: the pattern is increasing over time even if we look at conditional tax rates, obtained after purging from the impact of the evolution of per capita transfers which are the main alternative form of financing for municipalities. [4]

Table 2: Business property tax rates by year and geographic area.

	Observed						
Year	North West	North East	Center	South	Islands	Total	Conditional
1998	5.609	5.572	6.191	5.615	5.088	5.430	5.750
1999	5.670	5.647	6.294	5.747	5.275	5.492	5.820
2000	5.787	5.746	6.438	5.923	5.384	5.606	5.930
2001	5.846	5.816	6.495	6.007	5.523	5.670	6.020
2002	5.959	5.954	6.580	6.121	5.819	5.755	6.120
2003	6.035	6.125	6.630	6.217	5.868	5.826	6.190
2004	6.114	6.190	6.655	6.323	5.996	5.884	6.240
2005	6.124	6.340	6.714	6.391	6.066	5.971	6.320
2006	6.197	6.394	6.754	6.434	6.209	6.029	6.390

If budget variables do not fully explain tax setting decisions of Italian municipalities, do politics have a role? The answer seems to be positive: if we compare the proportion of jurisdictions delivering tax increases with the incidence of local council elections in each year, as in Figure 1, we notice a kind of a negative correlation. Decisions to rise property tax rates seem to be generally postponed to years immediately following elections (such as 2000 and 2005). Our empirical analysis is meant to explore more deeply this public choice dimension of local taxation, in particular by looking at spatial interactions as a potential channel through which electoral processes become relevant.

3 An economic model of tax setting decisions

Our analytical framework builds on information asymmetries and on leviathan-type governments, two assumptions which are standard in the YC literature (see Salmon, 1987 and Besley and Case, 1995). Only

[4]Entries of conditional are the coefficients β from the regression $y = \gamma_1 x + \gamma_2 x^2 + \sum_{i=1998}^{2006} \beta_i 1(Year = i)$, where y is the local tax rate, x is the per capita transfer and $1(Year = i)$ is an year dummy.

politicians are aware of the true cost of providing public services and, on the base of this information, decide the amount of local expenditures (and taxes). On the other hand, citizens have the power to sanction/praise the behaviour of their representatives through democratic elections. In this framework we can develop an agency model of tax setting decisions, whose outcome arises from the interaction between a principal (the representative voter) and an agent (the local official).

The representative voter

The representative voter in jurisdiction i acts under a budget constraint, according to which an exogenous level of income y is allocated to private consumption c and to the payment of a lump-sum local tax t. The individual budget constraint is then:

$$y_i = c_i + t_i \tag{1}$$

Individual preferences are positively affected by c_i as well as by the amount of public services provision (q_i); they are also affected by a set of fixed characteristics X_i specific to jurisdiction i (such as being on the beachfront or any other topographical feature). Hence the representative voter's utility function may be written as:

$$U_i = U(c_i, q_i; X_i) \tag{2}$$

where as usual $\partial u/\partial c > 0$ and $\partial u'/\partial c < 0$ and $\partial u/\partial q > 0$ and $\partial u'/\partial q < 0$. Substituting for c_i into eq. 2 yields:

$$U_i = U(y_i - t_i, q_i; X_i) \tag{3}$$

Eq. 3 states that individual utility decreases as long as the ratio t_i/q_i increases. As usual, the choice of the t_i/q_i ratio is delegated to the local official.

The local official

We assume for simplicity that the level of q_i is fixed. Local official derives a positive utility from the amount of rent diverted from public spending (s_i or waste), but politicians may differ according to the value they place on s. Given k as the exogenous unit cost of providing public goods and assuming that local expenditures are financed only via taxes, the budget constraint faced by local official can be written as:

$$t_i = k\, q_i + s_i \tag{4}$$

According to eq. 4 waste depends positively on tax rates, once the level of public services is fixed.

Suppose there are two time periods, and after period 1 there is an election in which voters choose between the incumbent and a challenger. Local official is interested in the amount of rent she can divert in period 1 as well as in period 2, the latter depending on the probability of being re-elected F_i. Hence the policy-maker objective function can be represented by:

$$V_i = V\{\theta(s_i), \delta F_i\} \tag{5}$$

where θ describes the value placed on waste and δ is a discount factor, denoting the subjective inter-temporal rate of substitution between current and future waste. The local official chooses t_i so as to maximise 5 subject to the budget constraint 4. Note that the more taxes are raised, the more public resources can be diverted for private purposes. But higher taxes for a given level of q imply a utility loss for the voter and hence affect the incumbent's probability of re-election F_i. In eq. 5 the role of F_i is crucial, since it links the voter's decision with that of the policy-maker.

The voting rule

As anticipated above, citizens are interested in the level of taxes for a given amount of public services but they are not able to observe the features of public good's supply, in particular they can not observe the unitary cost k. The best way for voters to circumvent the agency problem originated by asymmetric information, and gauge the extent of waste in their jurisdiction, is to look at taxes in similar jurisdictions (i.e. jurisdictions with comparable levels of q with a similar set of fixed characteristics X). We can suppose that these comparisons yield a maximum level of t_i/q_i that can be delivered for jurisdiction i's government to remain in office. This is a sort of "price cap" to public services provision in i and depends on the ratio t_j/q_j observed in other jurisdictions. Hence the probability of re-election can be described by the following voting rule:

$$F_i = \Phi(U_i \geq U_j) = \Phi\{(y_i - t_i, q_i; X_i) \geq (j_i - t_j, q_j; X_j)\} = \Phi\{(t_i/q_i \leq t_j/q_j); X_i, X_j\} \tag{6}$$

It is important to stress that through 6 what happens in jurisdiction j affects the decision making process in jurisdiction i. A decrease in the ratio t_j/q_j would make the relative price of public goods cheaper in j than in i; this would force i's incumbent to lower t_i in order to preserve the chance of being re-elected. Similarly, an increase in the ratio t_j/q_j would be detrimental to citizens in jurisdiction i, since it would allow an increase in t_i for any given level of public services. Yardstick comparisons, through information spillovers, generate the propensity of mayors to herd behind each other when setting tax rates.

Equilibrium

We can rewrite the objective function for the local official by substituting 4 and 6 into 5. This yields to:

$$V_i = V\{\theta(t_i - k\,q_i), \delta\Phi(t_i/q_i \leq t_j/q_j; X_i, X_j)\} \tag{7}$$

Note that t has opposite effects on the two components of V: on one hand, a greater t increases the amount of rent extraction for a given level of q; on the other hand, increasing t reduces the chances of re-election, other things being equal.

The first order condition for the maximization of 7 gives the following tax reaction function:

$$\theta'(k\,q_i) = \delta\Phi'(t_i/q_i \leq t_j/q_j; X_i, X_j) \tag{8}$$

Equation 8 states that, given the tax rates set in all other jurisdictions $j \neq i$, the policymaker chooses t_i so as to equate the marginal benefit of a tax increase (left hand side) with its marginal political cost (right hand side). In other words, local officials are faced with a dynamic trade-off between obtaining more perks in the present through a tax increase and loosing votes and hence the possibility to extract rents in the future. The balance between the two terms of the trade-off is determined by the parameters of the model, i.e. by the value each policy-maker places on wasting resources today (θ), her propensity to postpone waste in the future, the coerciveness of electoral features of the electoral framework which may value attached to rent seeking in the future (δ) the bindingness of the threat of not being re-elected (Φ', which may in turn depend from political institutions).

Welfare outcome

It is important to stress that policy-makers may differ according to the values of the parameters in 8. For instance, the disciplining effect of elections could be weaker for mayors subject to term limits and hence Φ' could be lower for them. Furthermore, politicians can display different values of θ and δ according to their type: "good" politicians may have low values for θ (and relatively high δ), being more reluctant to apply unjustified tax increases; "greedy" politicians may exhibit higher values for θ and relatively lower values for δ, thus being more prone to tax increases. The circumstance that not all politicians behave in the same way is crucial in order to assess the welfare consequences of yardstick competition. This point can be made by means of a simple example. Consider a world with two governments, a "good" one and a "bad" one, and suppose that both start from an equal level of tax rates t_0 (such as to fully cover the provision of public goods, whose unitary cost is k_0). At time one each government sets a new level for t; suppose they can choose between a low and a high value of taxes, respectively t_1^H and t_1^L (with $t_1^H - t_0 = \epsilon$ and $t_0 - t_1^L = \epsilon$).

Also the unitary cost for public goods provision can assume only two values at time 1: k_1^H , with probability π_G (π_B for bad government), and k_1^L with probability $1 - \pi_G$ ($1 - \pi_B$ for bad government). Let us first consider the behaviour of "good" government. The good mayor places a very low value on rent diversion for private purposes (i.e. θ_G is very low, put it equal to 0). Thus her tax setting decision at time 1 will depend solely on the dynamics of k, i.e. the mayor will opt for t_1^H if the unitary cost of public goods amounts to k_1^H and t_1^L otherwise. The expected value E_G of the tax rate in period 1 can thus be written as:

$$E_G(t_1) = \pi_G \, t_1^H + (1 - \pi_G) \, t_1^L. \tag{9}$$

Let us turn to "bad" government. The mayor, in this case, attributes a high value on rent extraction (θ_B is high) and is relatively risk-averse (she is greedy, being more reluctant to exchange rents today with the probability of rents tomorrow, i.e. δ is low). When setting tax rates at period 1, the bad mayor observes the choice made by the good one. If the latter chose t_1^H then the bad mayor will herd and choose the same tax rate, no matter the level of k: would k be high, opting for high taxes preserves the actual level of rents; would k be low then setting high tax rates would allow the bad mayor to increase her rents without loosing votes. But if the good government has set t_1^L then the bad mayor will have to look for the dynamics of k in her own jurisdiction. If the cost for public services has increased, then she will opt for t_1^H since she is not willing to accept a reduction in her actual level of rents in order to preserve her possibility of re-election tomorrow. If the cost for public services is k_1^L she will instead opt for t_1^L , since mimicking the good mayor would not require a reduction in her actual level of rents. Hence we can write the expected value of the tax rate set in period 1 by bad government as:

$$E_B(t_1) = \pi_G\, t_1^H + (1-\pi_G)\pi_B\, t_1^H + (1-\pi_G)(1-\pi_B)\, t_1^L. \tag{10}$$

The expressions 9 and 10 allow us to write down the (expected) average tax rate of the whole economy at period 1, call it t_1:

$$t_1 = \frac{E_G(t_1) + E_B(t_1)}{2} = \Psi t_1^H + (1-\Psi)t_1^L \tag{11}$$

with

$$\Psi = (2\,\pi_G + \pi_B - \pi_G\,\pi_B)/2 \tag{12}$$

It is easy to show that, under the sufficient condition that $\pi_G \geq 1/2$, it is always true that $t_1 > t_0$. This simple example is useful to show that the outcome of information spillovers in terms of social welfare is undetermined. In some cases, performance comparisons work as an alibi for bad incumbents to apply unjustified tax increases without any loss of votes, and then produce a "race to the top" rather than a "race to the bottom" among tax rates. A similar argument was put more formally by Besley and Smart (2002), who argued that tax competition can produce welfare adverse effects (i.e. higher levels of taxation for any given level of public services) when the fraction of bad types in the population of politicians is sufficiently high. Within this line of reasoning it is also important to stress the role for Φ i.e. of all the factors regulating the disciplining pressure exerted by voting on local tax setting behaviour. We can imagine, for example, that features such as a binding term limit (or a lower degree of tax salience) affect the intertemporal rate of substitution between actual and future waste, by reducing the expected loss of any tax increase. This may induce a higher fraction of bad politicians to mimick justified tax increases of good incumbents, thus leading to an increasing pattern in the average tax rate of the whole economy.

4 Empirical strategy

The natural empirical counterpart to the economic model set up in Section 3 is

$$t_{i,t} = \alpha_i + \beta \mathbf{X}_{i,t} + \rho \sum_{j \neq i} w_{ij} t_j + u_{i,t} \tag{13}$$

where the tax rate set by municipality i at time t is a function of a jurisdiction fixed effect, a vector $\mathbf{X}$ of observable characteristics and a weighted average of tax rates set by other municipalities. The last term is that the representative voter does not have full information regarding the budget fundamentals which may justify tax increases (such as a shock to the tax base or to the cost of providing public goods), thus a rational behaviour consists in making performance comparisons.In order to maximize their inter-temporal utility function, local officials include tax policies set by comparable jurisdictions in their reaction function.

Eq. 13 can be rewritten in matrix notation, stacking the observations by time and space, as:

$$\begin{aligned} t &= \mathbf{W}t\rho + \mathbf{X}\beta + u \\ u &= \lambda\mathbf{M}u + \epsilon \end{aligned} \tag{14}$$

where W and M are $N \times N$ weighting matrices, for example delimiting the geographical boundaries of spatial comparisons and are such that only their off-diagonal elements are non null and equal to ω_{ij}, $\mathbf{X}$ is a $N \times K$ matrix of covariates and ϵ and u are a well-behaved normal error. ρ and λ are spatial coefficients and β is a k vector of coefficients for the adjusting covariates.

To be consistent with our economic model, the empirical strategy is built upon three specifications, which are special cases obtained by imposing specific restrictions to the previous general model in 14.

First, we set $\rho = \lambda = 0$ and estimate a pure "autarchy" specification in which tax rates depend only on own municipal characteristics. This leads to our Model 1, defined as:

$$t = \mathbf{X}\beta + u. \tag{15}$$

Although simple and useful for many purposes, model 15 will be mis-specified as soon as there is (some) form of spatial correlation across municipalities, as suggested in Section 3. Thus, the second step consists in relaxing the hypothesis of absence of spatial interactions and consider the case in which the correlation enters *via* the conditional mean (i.e. $\rho \neq 0$ and λ=0). This produces our Model 2, which is a typical spatial autoregressive model (SAR):

$$t = \mathbf{W}t\rho + \mathbf{X}\beta + u. \tag{16}$$

According to eq. 16 tax rates in one municipality are jointly determined with that of neighbouring jurisdictions, and conditional on X if $\rho > 0$, own tax rates are high if neighbours' tax rates are high, irrespective of the reason why the latter are high. It is worth emphasizing that Model 2 is the empirical counterpart to YC:

in an attempt to fill the information gap, voters look at surrounding jurisdictions to infer the appropriate tax rate in their own municipalitites; all the relevant information for performance comparisons by voters is only concerned with observable tax rates.

Somewhat in between Eq. 15 and Eq. 16, our Model 3 is obtained by assuming that spatial correlation affects the structure of the disturbance term (i.e. $\rho = 0$ and $\lambda \neq 0$) and this gives rise to the following spatial error model (SEM):

$$t = \mathbf{X}\beta + u \tag{17}$$

$$u = \lambda \mathbf{M} t + \epsilon \tag{18}$$

According to eq. 17 and eq. 18 only the unexplained component of the tax rates tend to be spatially correlated. The economic interpretation is completely different from Model 2: here voters have a reasonable knowledge of the determinants of tax rates in own as well as in neighbouring jurisdictions (Bordignon et al, 2003). However, one of the drawbacks of the model is that empirically SEM can be a symptom of significant determinants of tax setting processes omitted from the model specification and spatially correlated (Elhorst, 2010 and Anselin et Al., 2006), making hard to draw firm conclusions from the results.

Also our estimating procedure is articulated in two distinct phases. In a first stage we stick to a pooled structure and ignore the time dimension of the data. This implies estimating our Model 1 through conventional OLS, while for Model 2 and 3 we resort to a Full Information Maximum Likelihood methodology to cope with the endogeneity arising from the fact that own and neighbours' tax rates are determined simultaneously (Anselin, 1998; for further details see also the Appendix).

At a second stage we switch to a panel approach, in order to cope with the risk of misspecification due to omitted variables and to overcome the bias arising in presence of individual time-invariant characteristics correlated with the observable covariates. In particular, when we consider the two specifications of Model 2 and Model 3 the panel approach may be useful to overcome two well known econometric issues in the spatial econometrics literature: possible spatial error dependence and possible correlation between the $\mathbf{X}$s and the error terms (see Brueckner, 2003). Spatial error dependence arises when the vector of error terms includes omitted variables that are themselves spatially dependent: for instance, tax rates may be used to finance public services related to topography or to forms of territorial associations among municipalities; ignoring these features, which are spatially correlated, would provide false evidence of strategic interactions in tax setting decisions. Correlation between the jurisdiction explanatory variables $\mathbf{X}$ and the error term may be due to characteristics unobservable to the researcher, such as local preferences: for instance, there could be a sorting of high income households in municipalities with specific topographical features and this would lead to an inconsistent estimate of the income coefficient, while also potentially distorting the estimates of the remaining coefficients. A panel data approach helps to tackle these problems, by representing all time-invariant community characteristics, either observed or unobserved, by community specific intercepts. These intercepts can then be treated as a vector of parameters to be estimated (fixed effect technique, FE), or as the realization of a latent random variable (random effect technique, RE). As we will see in Section 6, we will follow the former methodology and seek to explain the determinants of the individual intercepts by using the correction for the standard errors as suggested in Greene (2000).

5 The data set

The huge cross-section dimension of Italian municipalities, coupled with a great variety of the institutions responsible for the disclosure of statistics at municipal level made the construction of our data set extraordinarily burdensome. Variables are described in Table 3, which reports labels, measurement units, descriptive statistics and sources. We cover the 1998-2006 time span and, after dropping missing covariates, we end with a balanced panel of 6,123 yearly observations.

Our dependent variable is the business property tax rate, as anticipated in section 2. From the point of view of YC theory this could pose a problem, as the tax is paid by corporations or by owners of investment or vacation houses, who are often resident in a different jurisdiction from that levying the tax. But we agree with the argument made by Bordignon et al. (2003) that taxpayers are usually voters in the same municipality which is responsible for tax setting decisions, since the economic structure of Italian towns is dominated by small business whereas vacation houses are very concentrated (and account for a very limited share of ICI revenue).

Our matrix of explanatory variables, $\mathbf{X}$, is a set of local characteristics that we can distinguish in two subgroups: (a) standard covariates, i.e. traditional internal determinants of local tax setting policies, and (b) political variables.

The first subgroup is standard in all empirical analysis of tax setting decisions. This subgroup includes variables capturing peculiarities in public goods provision, such as scale economies or congestion effects (*Area*, *Pop*, *Urb*), and natural topographical features (*Alt*, *Beach*). We also take into account characteristics that may help determine preferences and local budget constraints by including in our regressions, respectively, the age structure of population (*Young*, *Old*) and two proxies for the availability of alternative sources of funding

(Y, G). Finally, a dummy variable discriminates municipalities laying in special statute regions (*Spec*), which have been devolved some peculiar functions by intermediates layers of government.[5]

Table 3: Summary Statistics

Variable	Def.	Mean	Std	10th	Median	90th	Obs.	Source
t_{bus}	Business property tax rate	5.7	0.8	4.5	6.0	7.0	71340	IFEL
t_{res}	Residential property tax rate	5.2	0.7	4.0	5.0	6.0	71079	IFEL
Area	Municipal area	0.4	0.5	0.1	0.2	0.8	72842	ISTAT
Alt	Height w.r.t. sea level	3.6	3.0	0.3	2.9	7.7	72842	ISTAT
Beach	Coastal municipialities	0.1	0.3	0.0	0.0	0.0	72842	ISTAT
Pop	Population	7.1	39.7	0.5	2.4	13.0	72842	ISTAT
Urb	Population per unit area	306.2	730.6	22.5	108.4	693.0	72796	ISTAT
Young	Share of population under 10 years	9.7	2.3	6.9	9.7	12.5	72841	Home Office
Old	Share of population over 65 years	21.4	6.5	14.3	20.6	29.7	72841	Treasury Dept.
G	Grants per capita	0.4	0.3	0.2	0.3	0.6	70245	Own calculation
Y	Income per capita	8.3	2.9	4.4	8.5	11.9	71336	
$Prop_{bus}$	Business property tax base	147.2	1201.4	7.3	37.9	245.2	71079	Home Office
Spec	Municipalities in special statute regions	0.2	0.4	0.0	0.0	1.0	72842	Own calculation
Elec	Election year	0.2	0.4	0.0	0.0	1.0	71481	Own calculation
Cycle	Days to election	40.7	16.7	17.4	41.4	63.1	71481	Own calculation
Comp	Strength of cabinet	0.5	0.2	0.3	0.6	1.0	71481	Own calculation
Reeleg	Second term mayors	0.5	0.5	0.0	0.0	1.0	71481	Own calculation
Civic	Civic coalition government	0.6	0.5	0.0	1.0	1.0	72842	Own calculation
Right	Right wing government	0.1	0.3	0.0	0.0	1.0	72842	Own calculation
NorthL	Northern league government	0.0	0.1	0.0	0.0	0.0	72842	Own calculation
Cent	Center government	0.1	0.2	0.0	0.0	0.0	72842	Home Office
Left	Left wing	0.3	0.4	0.0	0.0	1.0	72842	Home Office
Age	Age of the mayor	49.6	9.7	37.0	49.0	62.0	70337	ISTAT
Fem	Sex of the mayor	0.1	0.3	0.0	0.0	0.0	72842	ISTAT

The second subgroup of explanatory variables describes the political dimension of local budgeting behaviour, in the spirit of the public choice literature. As a matter of fact, tax setting decisions may be driven by incumbent's attempt to be re-elected. Our regressions hence take into account the electoral cycle by including a dummy for electoral years (*Elec*) together with a variable counting the number of missing days to the following elections (*Cycle*); computing the number of days to following elections can be an useful way to cope with misalignments between years in which tax choices are made, years in which they become effective and electoral deadlines.[6] We also consider some factors that may influence local fiscal discipline:

[5]Five special statute regions (Sicilia, Sardegna, Valle d'Aosta, Friuli-Venezia Giulia and Trentino Alto Adige) and two autonomous provinces (Trento and Bolzano) are entitled with major spending responsibilities and have access to specific funding from central government, under the form of tax shares over the main national taxes. Some of these functions are further devolved at the municipal level and this is reflected in municipalities' balance sheets, characterized by a greater availability of public funds (regional transfers in addition to State ones)

[6]Tax choices relative to year t are generally made between October of year *t-1* and April of year t. Thus electoral years and

the strength of local councils (*Comp*), which is computed as an Herfindal index of the share of each party figuring in local councils,[7] the absence of binding term limits (*Reelegib*) [8], the absence of explicit party affiliation (*Civic*), ideology i.e. mayors belonging to right wing, northern-league, center, left wing parties (*Right*, *Northl*, *Cent*, *Left*). The set of political explanatory variables is finally completed by the insertion of some mayor characteristics (*Age* and *Female*).

6 Estimation results

6.1 Model 1: the autarchy specification

The first set of results describes the "autarchy model" outlined in eq. 15, estimated through an OLS with standard errors robust to heteroskedasticity and clustering; estimates include a set of year and geographical dummies. Results are displayed in Table 4.

The most parsimonious specification in col. 1 includes only standard covariates, which prove to be mostly significant and in line with previous evidence. Demographic and structural variables play a role in tax setting decisions. In particular, tax rates are increasing with *Area* and *Urb*, suggesting that congestion inflates the costs public goods provision. Conversely, we found a negative impact for *Young* and *Old* (as in Bordignon et al., 2003). As for fiscal variables, we find that both Y and G tend to increase property tax rates: the first effect is consistent with public goods being normal goods, while the latter could be in principle associated with a *super-flypaper* effect of the type described by Gamkhar and Oates (1996), although this calls for further investigation.[9] Finally, municipalities located in special statute regions seem to suffer from less binding budget constraints and hence display lower levels of taxation. For the interpretation of the coefficients attached to political variables in col 2 it has to be borne in mind that the intercept refers to a municipality in Center Italy, led by a mayor who belongs to a civic coalition, in a non-electoral year and is under a binding term limit. The results show that politics and local tax setting decisions are deeply entrenched. The electoral cycle is relevant: being in an electoral year reduces tax rates almost by 8 percentage points

tax setting decisions do not necessarily overlap, as pointed out by Bordignon and Piazza (2010).

[7] The index is computed as: $I = \sum_p^n Q_p^2$ where Q_p is the fraction of representatives belonging to party p.The index ranges from $1/n$, when the seats in the council are equally divided among the n parties (maximum political fragmentation), to a value of 1, when only one party is represented (minimum political fragmentation)

[8] In Italy mayors cannot seat for two consecutive mandates, unless an early termination occurs (due to reasons different than voluntary resignation) before half term had expired.

[9] The theory would predict a crowding-out effect of grants with respect to tax revenues. In the framework of the *flypaper* literature, losses in transfers are usually compensated with tax increases by governments willing to preserve local expenditures ("fiscal replacement" effect observed by Gramlich, 1987). But Gamkhar and Oates (1996) found that local governments may magnify the spending response to cuts in grants by lowering own revenues as well: this gives rise to the "fiscal restraint" type of asymmetry also called *super-flypaper effect*.

and coherently tax rates tend to be higher the longer is the time span to following elections (*Elec*). The strength of political leadership (*Comp*) has a negative effect on tax rates, in line with the argument of Borge et al. (2008) that weak governments find it more difficult to resist pressures from local interest groups and thus are more prone to budget maximization. Also re-elegibility (*Reeleg*) has the expected negative impact on tax setting decisions and tax rates are even lower when the dummy is interacted with the election year: incumbents running for re-election apply tax rates lower by almost 12 per cent in electoral years. Both right and left wing parties levy higher taxes, whereas center parties and Northern League levy lower tax rates. This suggests that municipalities led by mayors subject to well defined party structures pay more attention to local public budgets, likely because they have a reputation to be maintained or because they wish to carry on their political career at the national level. Finally also the *age* of mayor has a reducing (although small) impact, as well as being a female (although the coefficient is less statistically significant).

However, pooling estimation may suffer from mis-specification either due to the intrinsic model or to data availability. Both drawbacks can be (partially) circumvented by exploiting the longitudinal dimension of our data, as we do in cols 3 and 4 of Table 4 where results for FE are displayed.[10] With respect to standard covariates, we can observe that the only relevant difference between cols 1 –3 and 2 – 4 relates to the sign of *Young* and *Old*, which now becomes positive: this indeed seems more consistent with the fact that population under 10 and over 65 are the targets of most of the services disclosed by municipalities (i.e. primary schooling and social assistance). As for political covariates, accounting for the panel dimension somehow confirms the robustness of the electoral cycle since both *electoral year* and *days to election* keep their explanatory power. The main differences concern the strength of local councils and the absence of term limits: the former becomes statistically insignificant for a technical reason (since *comp* is relatively constant over time the de-meaned variable is plenty of zeros by construction), whereas differences between lame ducks and re-elegible incumbents appear to be relevant only in electoral years. As for ideology, left wing and right wing parties still appear to levy higher taxes with respect to civic coalitions, whereas the affiliation to Northern League or to Center becomes ininfluent. Finally, the age of the mayor confirms to be negatively associated with tax rates, while sex is no longer relevant.

[10] The Hausman test strongly rejects the null hypothesis of equal coefficients across FE and RE (however RE estimates are available upon request from the authors). The advantage is that FE is consistent under less stringent conditions than RE.

6.2 Model 2: the SAR specification

Our subsequent step is to include a spatial structure in the empirical model. We do this by allowing correlation in tax rates, according to the specification of eq. 16: within this framework, a test for the null hypothesis that the reaction function's slope is zero is effectively a test for the existence of spillovers. The weighting matrix W is specified a priori and delimits the boundaries within which interactions take place; in particular, each element ω_{ij} indicates the relevance of other jurisdiction j in the process of interaction. As is customary in the literature, our choice is to adopt weights describing the relative location of municipalities based on a contiguity scheme under which ω_{ij} =1 for jurisdictions j sharing a border with i and ω_{ij} =0 in all other cases. The weights are row-standardised so that W_t corresponds to a vector of weighted averages of tax rates set by adjacent municipalities. We are reasonably confident that this scheme is a pragmatic representation of how information spreads out across municipalities, since citizens in contiguous jurisdictions are more likely to establish direct interactions (due to working or consumption paths) and neighbours belong to the same local media market. The main problem in estimating Model 2 is that tax rates are jointly determined by adjacent municipalities and hence the linear combination of the t_js appearing on the right hand side is endogenous and correlated with the error term ϵ_i. This means that OLS estimates are inconsistent, requiring use of other methodologies. As anticipated in Section 4, we estimated the reduced form of eq. 16 using maximum likelihood and a nonlinear optimization routine (since the key parameter ρ enters non linearly into the equation). A similar approach was adopted, among others, by Case, Hines and Rosen (1993), Besley and Case (1995), Bivand and Szymanski (1997), Brueckner and Saavedra (2001). Results are displayed in Table 5.

Let us first consider the pooled approach. As we can see by comparing columns 1 and 2 of Table 5 with the correspondent columns in Table 4, ML estimation of Model 2 yields a spatial coefficient for ρ of 0.43, which is statistically strongly significant: this means that roughly 40 per cent of tax setting decisions in one jurisdiction reflects a kind of herding with respect to policy choices made by adjacent municipalities. Overall the impact of standard covariates is confirmed, the main noticeable difference being the reduced magnitude of geographical dummies that reveals that roughly half of the clustering identified in Model 1 was indeed due to spatial contiguity. As for political variables, we can observe that spatial interaction absorbs part of the explanatory power of some political variables, in particular the effect of *re-elegib* and ideology. This seems to suggest that the influence of such variables is partly channelled through yardstick comparisons and ignoring the spatial correlation between municipalities would lead to an upward bias in the coefficients attached to political variables. [11]

Let us now turn to a FE panel approach, as described in columns 3 and 4 of Table 5.The key point to be emphasized is that spatial correlation is persistent, but now only 12 per cent of tax setting decisions reflect herding behaviour. With respect to the pooled approach roughly two thirds of spatial interactions are attributable to common time-invariant characteristics between contiguous jurisdictions. The relevance of political processes in tax setting decisions is substantially robust. In particular, the behavioural differences between re-eligible incumbents and lame ducks are somehow enhanced when we consider the panel dimension along with spatial interactions. It is important to stress that Bordignon et al.(2003) consider this behavioural differences are a pre-condition for YC to take place. Finally, not only age but also sex appear to be influent in the choice of property tax rates.

As a further exercise, to understand what is the nature of individual effects, we estimated the vector of the municipal intercepts using the correction for the standard errors suggested in Greene (2003) and Elhorst (2010) in order to cope with the fact that coefficients are unbiased but inconsistent. As we can see from Table 7 the fixed component of the tax rate is substantially explained by topographical features such as altitude, seaside or mountainside, or by area and institutional features.

6.3 Model 3: the SEM specification

As an ultimate proof for the robustness of our results we estimate the specification set by Model 3, as described in eq. 17, where the spatial structure enters *via* the error term. According to this specification only the components of the tax rates that are not explained by the **X**s tend to be correlated: this entails that voters have a reasonable knowledge of the economic fundamentals of tax setting decisions in own and neighbouring jurisdictions. Bordignon et al. (2003) consider this a reasonable assumption in the Italian context, where municipalities are indeed very small. Results are displayed in Table 6.

Overall, spatial correlation in municipal tax rates reaction function is confirmed. In the pooling specification the value for ρ is 0.44, a magnitude comparable to that found by Bordignon et al. (2003) for a sample of 143 municipalities. But the strength of spatial interactions could be over-estimated in presence of omitted variables that are themselves spatially dependent. Thus it is important to switch to a panel data approach, in order to wipe out all time-invariant community characteristics that may bias the estimate for ρ. Results in column 3–4 show that, as in the previous case, individual fixed effects explain roughly two thirds of spatial interactions. The impact of political variables is substantially confirmed: as in Model 2, being in

[11] Estimating Model 1 when the true model is Model 2 leads to an omitted variable bias equal to $\rho(\tilde{X}'\tilde{X})^{-1}(\tilde{X}'WY)$.

an electoral year reduces property tax rates by about 7 per cent, re-eligible mayors tend to apply tax rates lower by 2 per cent with an extra reducing impact arising from the interaction with elections

6.4 Robustness checks

The above results are coherent with the theoretical model in Section 3, but they must be checked against possible falsification tests. First of all, there might be some substitutability between the property tax rate and other forms of taxation, like for example the additional personal income tax or TARSU, and it is also important to understand the role played by the Residential tax rate. A second reason of concern, that has already been mentioned, is that, common to all the existing literature, the point estimates can be dramatically affected by the weighting matrix, thus different matrices should be investigated. Finally, the remarkable heterogeneity (broadly defined) of Italian municipalities requires a further understanding: in particular one may wonder whether results are valid for the universe of Italian municipalities or only for a smaller subset, and of course which subset.

In Table 9 we check the robustness of our results against these possible shortcomings. For the sake of brevity, for each check, we report only the coefficient for the spatial correlation (all the estimates are available upon request from the authors). The benchmark against which to compare the results is Table 5.

Not shown, we have also investigated various model specifications without dramatic difference in the estimated coefficients: such proofs include, but are not limited to, the inclusion of regional specific time dummies (against shocks spread over a single region, like for example regional specific budget constraint, or natural disasters), different definitions of time elapsed to the next elections, and using different indicators for the population structure (using the dependency ratio rather than the share of young and old people).

Business Property Tax rate vs Other forms of taxation

In principle, the property tax rates can be substitute or complement to the additional personal income tax and TARSU. In fact, the former was kept at a constant (municipal) rate by the Italian Central Government during the period under study, thus there is no municipal autonomy in it, the latter is a tariff rather than a tax, covering a given share of waste collection costs (see Section 2): as a consequence, we do not pursue these checks here.

Most important is the interplay between the business and the residential property tax rate. Whilst for the reasons in Section 2 we focus exclusively on the *business tax rate*, one might wonder whether results are robust using the *residential tax rate*. The row *Residential Tax* in Table 9 is reassuring that the economic model set up in Section 3 is valid for the tax rates where Italian municipalities have autonomy. Indeed, the estimated spatial coefficient is close to (actually, slightly smaller than) the coefficient estimated using the business property tax rate. Most important we can propose a justification for a puzzling result, which is routinely encountered in this literature. Looking at what neighbors do should bring to a race to the bottom in the tax rates, whereas with the business tax rate we see a race to the top; using the residential property tax rate there is a constant across an average level: the puzzling result can be reconciled looking deeper into the function Φ. **qui integrare alla luce di modello economico**

Different weighting matrix

A major concern with spatially correlated data is that there exist infinite possible weighting matrices, each capturing a different legitimate definition of "proximity". In this paper we focus exclusively on geographical distances, but we also estimate a separate model by population size, which can be viewed as jointly considering the geographical definition and the population definition of proximity. So far we have used what is probably the simplest definition of proximity, based on nearest contiguous neighborhood (this is our benchmark throughout this section). Nevertheless, because at the bottom of the yardstick competition theory is the flow of information across jurisdictions, such definition might be too restrictive. In this vein, one would ideally use a definition based on local media network (see Revelli, 2007 and 2008, for the UK): this detail is unavailable to us, thus we simply extend the definition of the neighborhood beyond the nearest, under the rational that near jurisdictions have the same local media network and that, within a given media network, voters are more concerned with the outcome in nearest jurisdictions than those that are further away. The same approach of using different *geographic* matrices is followed in other studies like for example Basley and Case (1995), Brueckner and Saavedra (2001), Revelli (2003) to mention some.

As definition of neighborhood we set the 5 and the 10 nearest neighbors. The numbers were chosen for two reasons: the first is an institutional reason that unions of municipalities are typically built up of small numbers of municipalities (generally smaller than than 10, on average 5); the second, in the spirit of the "network matrix" in Revelli (2007 and 2008), is that inhabitants reading local newspapers are interested in their "local market", which we conventionally set to a maximum of 10. Of course one can play around with these number, but we argue below that the model is largely robust to this decision.

As is clear from the first two rows in Table 9 the estimated coefficients for the spatial correlation coefficients are very similar to those in Table 5. In particular, using 5 nearest neighbors the spatial correlation coefficients are close to those in Table 5, whereas using 10 neighbors the SAR coefficients are always slightly higher, no matter the method that we use.[12] The relevant result is that in both cases as going from pooled models to habit formation models, the spatial correlation is positive and largely significant, but under our economic model of habit formation the estimated spatial correlation coefficient is as half as that of the pooled model. These robustness checks are indicative that, once the yardstick competition is considered, the weighting matrix that is used can change the point estimates, but, as far as matrices are appropriate, the theory is confirmed.

Homogenous population sizes

Another reason of concern is related to the different population sizes. In Section 2 we emphasized that Italian municipalities are the ideal laboratory to test the yardstick competition theory because, as municipalities are the lowest government level, mayors are subject to stronger reputational pressures. More precisely, it can be argued that the smaller is the municipality the stronger are the reputational pressures and thus the closer to neighbor jurisdictions are the tax rates, in which case we expect the spatial coefficients to be larger than in the benchmark case. To test this argument we have split the sample over two dimensions: one is legal and distinguishes between Administrative Capitals (106 municipalities) vs. NON-Administrative Capitals (about 8,000 municipalities), the other distinguishes between sizes greater than 5,000 inhabitants (about 2,200 municipalities, or 72% of the total sample) or smaller than that (slightly less than 5,900 municipalities, about 28% of the sample). Municipalities whose population size is above the threshold are subject to the Internal Stability Pact.

Of course, Administrative Capitals do not have contiguous neighbors that are themselves Administrative Capitals, thus the analysis is only on NON-Administrative Capitals. For NON-Administrative Capitals, that are the greatest majority of our sample, all the estimated coefficients are undistinguishable from those estimated for the whole sample (as expected). Similarly, for municipalities larger than 5,000 inhabitants the SAR coefficient is lower than the benchmark with all the models. As soon as we move to the smallest municipalities, the coefficients are slightly higher than those estimated using the whole sample, as predicted from the reputational argument of the yardstick competition theory.

Hence these robustness checks are reassuring that *i*) the yardstick competition is a valid theory for all

[12] The statistical reason for the coefficient increasing in the number of neighbours is that the influence function of the mean is infinite (Jurečová and Picek, 2006) and, along with the stylized fact of increasing rate over the period and a positive spatial correlation, enlarging the neighbourhood more likely to find higher tax rates in nearby jurisdictions, hence the estimated coefficient slightly larger.

the municipalities, but *ii*) for smallest municipalities its pressure on the mayors is stronger (i.e., the spatial correlation coefficient is higher).

7 What drives herding behaviour?

The econometric analysis in the previous section showed that municipal property tax rates are spatially linked. The panel approach enabled us to purge out similarities due to spatial autocorrelation of omitted variables and to take account of the effect of unobserved time invariant characteristics. Hence we are rather confident that our spatial coefficient (about 0.10-0.12), which is sensibly lower than previous cross-section evidence, captures substantial strategic interaction do to herding behaviour.[13] Another key result is that the impact of neighbouring taxing policies (as estimated in Model 2) seems to be stronger than that of traditional (as well as political) determinants of tax setting decisions.

But the evidence on herding behaviour nothing says about its underlying determinants. Indeed the general cathegorization provided by Brueckner (2003) highlights that informational spillover and resource flows models are observationally equivalent since they both generate reaction functions in which the decision variable for a given jurisdiction depends on the choices of other jurisdictions. Hence the spatial coefficient ρ may reflect information spillovers in the framework of electoral competition, as stated by the YC theory, as well as municipalities competing with each other in order to trigger resource flows, as stated by the standard tax competition theory (STC hereby). The question is how to ascertain what is the true nature of spatial correlation in property tax rates.

Our answer is that the strategic interaction that we observe in Italian property tax rates is driven by information spillovers. This opinion is essentially based on two counterfactual arguments. First, STC would imply that local tax setting decisions are effectively able to influence location preferences of firms and households and we can rule out that this happens in the case of property taxes, since moving the tax base would require transaction costs which are very likely to offset any possible gain in tax rates. This is all the more true in the Italian framework, where local heterogeneities in labour market conditions, cost of living, quality of public services are sizeable and are consequently more relevant in influencing location choices of economic agents. The second counterfactual argument is that, if strategic interactions were driven by resource flows, then we would observe a typical "race to the bottom" pattern in tax rates, which in fact

[13] Some of the problems due to spatial error dependence or to correlation between the error terms and the explanatory variables may survive to first differencing if they arise as a consequence of time-varying community characteristics. However the extent of these problems is expected to be very limited and could be partly mitigated by the inclusion of time dummies.

never occurred to date (as already pointed in Section 5). The outcome of increasing tax rates would instead be compatible with information spillovers triggering some form of tax collusion in the framework of a high proportion of politicians of the rent seeking type (as formally pointed out by Besley and Smart, 2002).

We are aware assessing ultimately the source of tax herding additional would require additional evidence. In particular, further efforts should be addressed to the estimation of the structural equations generating tax reaction functions. One line of research could be estimating an auxiliary equation relating incumbent chances of re-election to taxes in neighbouring jurisdictions (as in Besley and Case, 1995) or, alternatively, exploring the nexus between a jurisdiction tax base and its tax rate (as in Brett and Pinske, 2000). Another interesting line of research could be to try to assess empirically the quality of local politicians and explore its relation with herding and tax increases. These empirical issues are at the core of the agenda for future developments of our analysis.

8 Conclusions

The issue of mutual interactions in local tax setting policies has been attracting a growing theoretical and empirical research over the last few years. A specific explanation for spatial similarities is given by the yardstick competition theory, according to which strategic interactions are due to information spillovers between voters who look at public services and taxes in other jurisdictions to judge whether their government deserves to be re-elected or not.

Despite a growing emphasis on the spatial dimension of local fiscal policies, research on the Italian case is still scarce, the only reference work being that by Bordignon et al. (2003) to date. We seek to supplement this dearth of evidence and study spatial dependence in property tax rates set by Italian municipalities in the 1998-2006 period.

Our results support that spatial interactions among adjacent municipalities are positive and significant, revealing a propensity to herding in local tax setting policies. The discussion of the results has focused on two main questions: to what extent spatial correlation reflect genuine mimicking behaviour rather than error correlation due to model mis-specification? How can it be excluded that the true driving forces are resource flows in line with standard tax competition models?

To tackle with the first issue we exploited the panel dimension of our data set, so as to purge the effect of time-invariant observable and unobservable characteristics of each municipality that may induce spatial correlation in the error term. We observe that spatial interaction persists in panel specifications, although its magnitude almost reduces by two thirds. However the effect of neighbouring choices on own tax rates still appear more relevant than traditional tax setting determinants.

Our answer to the second issue is mostly counterfactual. We are confident that the source of this spatial correlation has to do with informational externalities since residential or business properties can not be considered as a mobile tax base. Furthermore, standard tax competition is not supported by the stylised evidence of increasing property tax rates over time (if municipalities were to compete in order to attract tax bases we would have observed a typical "race to the bottom" pattern in tax rates).

Finally, we believe that from this stylised evidence one can not exclude that tax herding is a symptom of tax collusion among local governments, as predicted by the model of Besley and Smart (2003) when bad type politicians are predominant. Hence it is crucial to set up policy measures aimed at increasing citizens' awareness on the economic fundamentals of local policy choices, on one hand, and at improving the quality of local politicians through electoral institutions that may improve ex post selection of good type incumbents.

References

Alesina, A. and Rosenthal, H. (1995), *Partisan Politics, Divided Governments, and the Economy*, Cambridge Univ. Press, New York.

Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Dorddrecht: Kluwer Academic Publishers.

Anselin L., Le Gallo J., Jayet H. (2006), "Spatial panel econometrics". In: Matyas L, Sevestre P. (eds) *The econometrics of panel data, fundamentals and recent developments in theory and practice*, 3rd edn. Kluwer, Dordrecht, pp 901-969.

Baltagi, B.H. (2001), *Econometric Analysis of Panel Data*, 2nd ed., John Wiley & Sons, Ltd: New York

Besley, T., Case, A. (1995a), "Incumbent behavior: Vote Seeking, Tax-Setting, and yardstick Competition,", *The American Economic Review*, 85: 25–45.

Besley, T., Case, A. (1995b), "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits", *The The Quarterly Journal of Economics*, 110: 769–798.

Besley, T., Smart, M. (2002), "Does Tax Competition Riase Voter Welfare", CEPR discussion paper n. 3131.

Bivand, R., Szymanski, S. (1997), "Spatial dependence through local yardstick competition: Theory and Testing",*Economic letters*, 55: 257–265.

Bordignon, M., Cerniglia, F. and Revelli, F. (2003), "In search of yardstick competition: a spatial analysis of Italian municipality property tax setting", *Journal of Urban Economics*, 54: 199–217.

Brett, C., Pinske, J. (2000), "The determinants of municipal tax rates in British Columbia", *Canadian Journal of Economics*, 33: 695–714.

Brueckner, J. K., Saavedra, L. A. (2001), "Do local governments engage in strategic tax competition?", *National Tax Journal*, 54: 203–229.

Brueckner, J. K.(2003), "Strategic Interaction Among Governments: an Overview of Empirical Studies", *International Regional Science Review*, 26: 175–188.

Case, A.C., Rosen, H.S. and Hines, J. (1993), "Budget spillovers and fiscal policy interdependence : Evidence from the states", *Journal of Public Economics*, 52: 285–307.

Elhorst, J.P. (2010), "Spatial Panel Data Models". In Fischer MM, Getis A (Eds.) *Handbook of Applied Spatial Analysis*, pp. 377-407. Springer: Berlin Heidelberg New York.

Elhorst, J.P. (2003), "Specification and Estimation of Spatial Panel Data Models". *International Regional Science Review*, 26: 244-268.

Gamkhar, S., Oates, W. (1996), "Asymmetries in the Response to Increases and Decreases Intergovernmental Grants: Some Empirical Findings", *National Tax Journal*, 49(4): 501-512.

Gramlich, E.M. (1987), "Federalism and Federal Deficit Reduction", *National Tax Journal*, 40(3): 299-313.

Greene, W.H. (2000), *Econometric Analysis*, Prentice Hall, New Jersey.

Heyndels, B. and Vuchelen, J. (1998) "Tax Mimicking Among Belgian Municipalities", *National Tax Journal*, 51: 89–101.

Kelejian, H. and Prucha, I. (1999) "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model", *International Economic Review*, 40: 509–533.

Lazear, E.P. and Rosen, S. (1981) "Rank-Order Tournaments as Optimum Labor Contracts", *Journal of Political Economy*, 89(5): 841–864.

LeSage, J.P. (1998) "*ECONOMETRICS: MATLAB toolbox of econometrics functions*", Statistical Software Components, Boston College Department of Economics

Manki, C. (1993) "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies*, 60: 431–442.

Nalebuff, B. and Stiglitz, J. (1983) "Prices and Incentives: Towards a General Theory of Compensation and Competition", *Bell Journal of Economics*, 14(1): 21–43.

Revelli, F. (2008) "Performance competition in local media markets", *Journal of Public Economics*, 92: 1585–1594.

Revelli, F. (2007) "Local media networks and social spending: Evidence from the UK", *Economics Letters*, 96: 144–149.

Revelli, F. (2005) "On Spatial Public Finance Empirics", *International Tax and Public Finance*, 12: 475–492.

Revelli, F. (2002) "Testing the Tax Mimicking versus Expenditure Spill-Over Hypotheses Using English Data", *Applied Economics*, 34: 1723–1731.

Revelli, F. (2001) "Spatial Patterns in Local Taxation: Tax Mimicking or Error Mimicking", *Applied Economics*, 33: 1101–1107.

Salmon, P. (1987) "Decentralisation as an Incentive Scheme", *Oxford Review of Economic Policy*, 3(2): 24–43.

Santolini, R. (2007) "La letteratura empirica sullinterdipendenza e sul comportamento imitativo nelle scelte di politica fiscale", *Studi e Note di Economia*, 12: 417–433. In Italian.

Shleifer, A. (1985) "A Theory of Yardstick Competition", *The RAND journal of Economics*, 16(3): 319–327.

Solé Ollé A. (2003) "Electoral accountability and tax mimicking: the effects of electoral margins, coalition government, and ideology", *European Journal of Political Economy*, 19: 685–713.

Table 4: Tax setting under autarchy specification (Model 1): 48,984 Observations, 6,123 Municipalities, 8 years.

Variable	Pooled		Fixed Effects		Random Effects	
	Standard 1	Political 2	Standard 3	Political 4	Standard 5	Political 6
Population	0.000 ***	0.000 ***	0.019 **	0.018 **	0.000	0.000
Urbanization	0.000 ***	0.000 ***	0.000	0.000	0.000 ***	0.000 ***
Share of young	-0.012 ***	-0.012 ***	0.017 ***	0.017 ***	0.014 ***	0.013 ***
Share of old	-0.002 **	-0.002 **	0.009 ***	0.008 ***	0.006 ***	0.005 ***
Grants per capita	-0.147 ***	-0.130 ***	-0.055 ***	-0.055 ***	-0.070 ***	-0.068 ***
Income per capita	0.017 ***	0.013 ***	0.007 **	0.006 **	0.010 ***	0.009 ***
Special Reg.	-1.071 ***	-1.038 ***			-1.121 ***	-1.112 ***
North West	-0.412 ***	-0.371 ***			-0.386 ***	-0.375 ***
North East	-0.215 ***	-0.193 ***			-0.190 ***	-0.185 ***
Southern	-0.515 ***	-0.510 ***			-0.552 ***	-0.553 ***
Islands	0.079 ***	0.072 ***			0.089 *	0.087 *
Area (Km sq)	0.068 ***	0.043 ***			0.070 ***	0.064 ***
Altitude	0.001	0.004 ***			-0.003	-0.003
Mountains	0.049 ***	0.053 ***			0.040 **	0.042 **
Beach	0.557 ***	0.557 ***			0.557 ***	0.556 ***
Right Wing		0.072 ***		0.039 ***		0.043 ***
Northern League		-0.088 ***		-0.001		-0.012
Left Wing		0.075 ***		0.030 ***		0.036 ***
Center		-0.052 ***		0.012		0.008
Election Year		-0.075 ***		-0.075 ***		-0.075 ***
Term limit		-0.118 ***		-0.012		-0.023 **
Term limit × Election		-0.039 ***		-0.010 *		-0.012 **
Strength cabinet		-0.204 ***		0.011		-0.007
Days to election		0.001 ***		0.001 ***		0.001 ***
Female		-0.014		-0.015		-0.015
Age		-0.008 ***		-0.004 **		-0.004 **
Age sq.		0.000 ***		0.000 **		0.000 **
Intercept	5.878 ***	6.127 ***			5.512 ***	5.603 ***
R^2 sq.	0.241	0.271	0.242	0.247	0.261	0.265
-Log Like	51892	51643	11742	11585	15790	15395

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.
Year dummies are included in the regression.

Table 5: Tax setting under SAR specification (Model 2): 48,984 Observations, 6,123 Municipalities, 8 years.

Variable	Pooled		Fixed Effects		Random Effects	
	Standard	Political	Standard	Political	Standard	Political
	1	2	3	4	5	6
Population	0.000 ***	0.000 ***	0.016 ***	0.015 ***	0.000	0.000
Urbanization	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000	0.000
Share of young	-0.008 ***	-0.007 ***	0.018 ***	0.018 ***	0.015 ***	0.014 ***
Share of old	0.000	0.000	0.007 ***	0.007 ***	0.006 **	0.006 **
Grants per capita	-0.108 ***	-0.094 ***	-0.070 ***	-0.070 ***	-0.079 ***	-0.078 ***
Income per capita	0.015 ***	0.013 ***	0.007 ***	0.007 ***	0.010	0.009
Special Reg.	-0.675 ***	-0.658 ***			-0.887 ***	-0.878 ***
North West	-0.246 ***	-0.211 ***			-0.313 ***	-0.303 ***
North East	-0.118 ***	-0.100 ***			-0.158 ***	-0.154 ***
Southern	-0.317 ***	-0.320 ***			-0.486 ***	-0.488 ***
Islands	0.074 ***	0.075 ***			0.021	0.021
Area (Km sq)	0.074 ***	0.053 ***			0.087	0.082
Altitude	0.007 ***	0.009 ***			0.003	0.004
Mountains	0.024 ***	0.025 ***			0.039	0.040
Beach	0.417 ***	0.420 ***			0.504 ***	0.505 ***
Right Wing		0.053 ***		0.022 ***		0.026 ***
Northern League		-0.070 ***		0.018		0.004
Left Wing		0.070 ***		0.023 ***		0.030
Center		-0.024 *		0.016 *		0.013 **
Election Year		-0.081 ***		-0.066 ***		-0.066 ***
Term limit		-0.064 ***		-0.020 **		-0.029 ***
Term limit × Election		-0.021		-0.018 ***		-0.020 **
Strength cabinet		-0.168 ***		-0.002		-0.016 ***
Days to election		0.001 **		0.000		0.000
Female		-0.009		-0.015 *		-0.014 ***
Age		-0.011 ***		-0.004 ***		-0.004
Age sq.		0.000 ***		0.000 ***		0.000
SAR	0.402 ***	0.399 ***	0.129 ***	0.125 ***	0.184 ***	0.183 ***
Intercept	3.429 ***	3.782 ***			4.393 ***	4.510 ***
R^2	0.327	0.3321	0.858	0.859	0.838	0.839
-Log Like	49716	49543	11583	11432	24903	24762

Note: *** is 1% significance level; ** is 5% significance level; * is 10% significance level.
Year dummies are included in the regression.

Table 6: Tax setting under SEM specification (Model 3): 48,984 Observations, 6,123 Municipalities, 8 years.

Variable	Pooled		Fixed Effects	
	Standard	Political	Standard	Political
	1	2	3	4
Population	0.000 ***	0.000 ***	0.015 ***	0.014 ***
Urbanization	0.000 ***	0.000 ***	0.000 ***	0.000 ***
Share of young	-0.007 ***	-0.006 **	0.018 ***	0.018 ***
Share of old	-0.001 *	-0.001	0.007 ***	0.007 ***
Grants per capita	-0.076 ***	-0.070 ***	-0.068 ***	-0.068 ***
Income per capita	0.021 ***	0.019 ***	0.007 ***	0.007 ***
Special Reg.	-1.109 ***	-1.094 ***		
North West	-0.396 ***	-0.360 ***		
North East	-0.206 ***	-0.187 ***		
Southern	-0.513 ***	-0.516 ***		
Islands	0.125 ***	0.127 ***		
Area (Km sq)	0.079 ***	0.057 ***		
Altitude	0.009 ***	0.010 ***		
Mountains	0.012	0.014		
Beach	0.472 ***	0.474 ***		
Right Wing		0.058 ***		0.022 ***
Northern League		-0.041 *		0.017
Left Wing		0.080 ***		0.023 ***
Center		-0.002		0.016 *
Election Year		-0.089 ***		-0.069 ***
Term limit		-0.031 **		-0.021 **
Term limit × Election		-0.017		-0.019 ***
Strength cabinet		-0.160 ***		-0.002
Days to election		0.001		0.000
Female		-0.005		-0.014 *
Age		-0.011 ***		-0.004
Age sq.		0.000 ***		0.000
SEM	0.422	0.409 ***	0.133 ***	0.128 ***
Intercept	5.745 ***	6.074 ***		
R^2	0.238	0.245	0.857	0.858
-Log Like	49758	49605	11591	11436

Note: *** is 1%; ** is 5%; * is 10% significance level.
Year dummies are included in the regression.

Table 7: Explaining individual characteristics. From "Political" model. P-values are based on S.E. corrected as in Elhorst (2010) and Greene (2000).

Variable	OLS	SAR	SEM
Initial Tax Base	-0.477 ***	-0.464 ***	-0.503 ***
Altitude	0.009 **	0.008 **	0.009 **
Beach	0.505 ***	0.554 ***	0.549 ***
Area (Km sq)	0.067 ***	0.072 ***	0.073 ***
Special Reg.	-0.908 ***	-1.021 ***	-1.023 ***
Mountains	0.083 ***	0.091 ***	0.100 ***
NW	-0.333 ***	-0.382 ***	-0.387 ***
NE	-0.168 ***	-0.197 ***	-0.199 ***
S	-0.563 ***	-0.637 ***	-0.641 ***
IS	-0.033	-0.055	-0.050
Intercept	4.693 ***	5.436 ***	5.381 ***

Note: *** is 1% significance level; ** is 5% significance level;
* is 10% significance level.

Table 8: Tax rates (actual and conditional on the effect of transfers from Central Government) and per capita transfers from Central Government.

Year	Observed	Conditional	Transfer
1998	5.43	5.75	313.6
1999	5.49	5.82	316.9
2000	5.61	5.93	313.4
2001	5.67	6.02	358.6
2002	5.75	6.12	370.2
2003	5.83	6.19	375.1
2004	5.88	6.24	365.6
2005	5.97	6.32	362.5
2006	6.03	6.39	366.2

Entries of conditional are the coefficients β from the regression $y = \gamma_1 x + \gamma_2 x^2 + \sum_{i=1998}^{2006} \beta_i 1(Year = i)$, where y is the local tax rate, x is the per capita transfer and $1(Year = i)$ is an year dummy.

Table 9: Some robustness checks: Entries are the coefficient of SAR for model "Political".

Robustness	Pool	FE	RE
Residential Tax	0.395 ***	0.119 ***	0.171 ***
5 Nearest	0.400 ***	0.136 ***	0.180 ***
10 Nearest	0.509 ***	0.198 ***	0.261 ***
NON Capital	0.401 ***	0.121 ***	0.182 ***
$> 5,000$	0.300 ***	0.100 ***	0.145 ***
$< 5,000$	0.410 ***	0.151 ***	0.203 ***

Figure 1: Municipalities: percentage of tax increases and elections by year

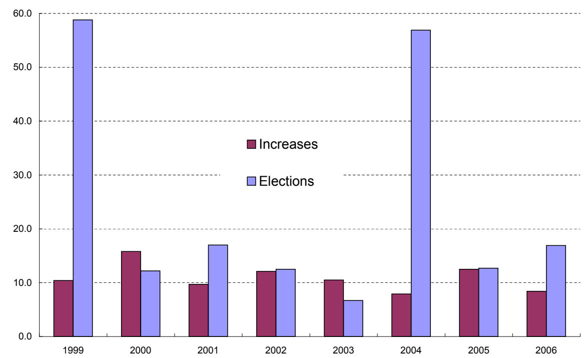
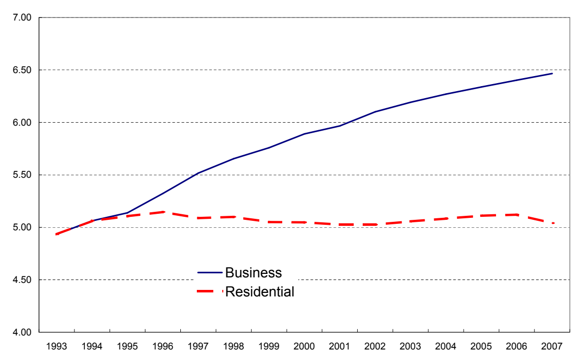


Figure 2: Business and Residential municipal tax rates by year



9 Appendix

Two different approaches exist to estimate a Spatial model as in eq. 16 or eq. 17– 18: one is through GMM (see Kelejian and Prucha, 1998, 1999 for further details), one is through a Full Information Maximum Likelihood (FIML; see Anselin, 1988 for further details): in this paper we have used the latter approach for reasons that can be seen in Revelli (2001).

Under normality assumption for the underlying distribution function, (log)Likelihood of SAR model (eq. 16) would be

$$\ln L(\sigma^2, \rho, \beta; y) = c(y) - \frac{N}{2}\ln(\sigma^2) + \ln|I - \rho W| - \frac{1}{2\sigma^2}[(I - \rho W)Y - X\beta]'[(I - \rho W)Y - X\beta],$$

for SEM (eq. 17– 18) would be

$$\ln L(\sigma^2, \rho, \beta; y) = c(y) - \frac{N}{2}\ln(\sigma^2) - \frac{1}{2}\ln|(I-\rho W)^{-1}(I-\rho W)^{-T}| - \frac{1}{2\sigma^2}(Y - X\beta)'[(I-\rho W)^{'}(I-\rho W)](Y-X\beta), \tag{19}$$

and useful results for the likelihood maximization exist to get rid of the terms $\ln|\cdot|$ (Anselin, 1988).

As we outline in Section 4, even the most sophisticated model can be incomplete. One source of incompleteness might be due to time invariant municipal characteristics: the non-zero correlation between the explanatory variables and the unobservable individual characteristics may induce a bias in the point estimates. Appropriately exploiting the longitudinal dimension of our data we can avoid cush bias.

To estimate a FE model, robust to such mispecifications, Elhorst (2010) suggests a data transformation identical to the classical FE transformation, such that the (log)likelihood function becomes

$$-\frac{NT}{2}\ln(2\pi\sigma^2) + T\sum_{i=1}^{N}\ln(1 - \rho w_i) - \frac{1}{2\sigma^2}\sum_{t=1}^{T} u_t^{'} u_t$$

where $u_t = (I - \rho W)\left[Y_t - \overline{Y_t} - (X_t - \overline{X_t})\beta\right]$ and $u_t = (I - \rho W)\left[Y_t - \overline{Y_t}\right] - (X_t - \overline{X_t})\beta$ for SEM and SAR, respectively, and where $\overline{Y_t}$ and $\overline{X_t}$ are averages over individuals. Following the approach for the classical case, it is also possible to estimate a RE model, although for SEM will be computationally cumbersome (Elhorst, 2003 and 2010).

The estimation proceeds using standard softwares.