ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Bordignon, Massimo; Cerniglia, Floriana; Revelli, Federico

Working Paper

In Search for Yardstick Competition: Property Tax Rates and Electoral Behavior in Italian Cities

CESifo Working Paper, No. 644

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Bordignon, Massimo; Cerniglia, Floriana; Revelli, Federico (2002) : In Search for Yardstick Competition: Property Tax Rates and Electoral Behavior in Italian Cities, CESifo Working Paper, No. 644, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/76046

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

A joint Initiative of Ludwig-Maximilians-Universität and Ifo Institute for Economic Research

Working Papers

IN SEARCH FOR YARDSTICK COMPETITION: PROPERTY TAX RATES AND ELECTORAL BEHAVIOR IN ITALIAN CITIES

Massimo Bordignon
Floriana Cerniglia
Federico Revelli

CESifo Working Paper No. 644 (1)

January 2002

Category 1: Public Finance

CESifo
Center for Economic Studies & Ifo Institute for Economic Research
Poschingerstr. 5, 81679 Munich, Germany
Phone: +49 (89) 9224-1410 - Fax: +49 (89) 9224-1409
e-mail: office@CESifo.de
ISSN 1617-9595

An electronic version of the paper may be downloaded
- from the SSRN website: www.SSRN.com
- from the CESifo website: www.CESifo.de

CESifo *Working Paper No. 644*
January 2002

IN SEARCH FOR YARDSTICK COMPETITION: PROPERTY TAX RATES AND ELECTORAL BEHAVIOR IN ITALIAN CITIES

Abstract

Do citizens engage in comparative performance evaluation across local governments? And if they do, how can we disentangle this behavior from other forms of strategic interactions among local governments or simple spatial correlation across neighboring jurisdictions? We use spatial econometrics techniques and the institutional characteristics of the Italian system to test if some theoretically derived predictions of yardstick competition theory are supported by data, estimating to this aim both a tax setting and a popularity equation. The results show that local tax rates are positively auto-correlated among neighboring jurisdictions when the mayors run for re-election, while this correlation is absent where either the mayors face a term limit or where they are backed by an overwhelming majority in the local council. Both results are in clear agreement with yardstick theory. On the other hand, the results of the estimation of the popularity equation are less supportive of the theory, possibly as a result of the difficulty in controlling for public service quality and the simultaneous setting of multiple policy instruments.

JEL Classification: C51, D72, H71.

Keywords: local property tax setting, yardstick competition, spatial auto-correlation.

Massimo Bordignon
Università Cattolica del Sacro Cuore
Facoltà Economia
L.go Gemelli 1
20123 Milano
Italy
Mbordig@mi.unicatt.it

Floriana Cerniglia
Università Cattolica del Sacro Cuore
Facoltà di Scienze Politiche
L.go Gemelli 1
20123 Milano
Italy

Federico Revelli
Università di Torino
Facoltà di Scienze Politiche
Via Po 53
10124 Torino
Italy

1. Introduction

One of the most interesting recent developments in empirical local public finance is the growing evidence of the existence of strategic interaction in fiscal behavior across local governments. Ladd (1992), Case (1993), Case *et al.* (1993), Kelejian and Robinson (1993), Murdoch *et al.* (1993), Besley and Case (1995a), Shroder (1995), Bivand and Szymanski (1997, 2000), Buettner (2001), Heyndels and Vuchelen (1998), Brueckner (1998), Brueckner and Saavedra (2001), Figlio *et al.* (1999), Brett and Pinkse (2000), Saavedra (2001), Revelli (2001a; 2001b) and Solè Ollè (2001) all found evidence of strategic behavior, using a variety of data sets, endogenous variables and estimation techniques. While the empirical evidence seems robust, unfortunately it is consistent, or it has been taken as such, with several competing theories of local government behavior. For example, the fact that local tax rates – or more generally fiscal variables – are correlated across neighboring jurisdictions has been variably interpreted as arising from simple mimicking behavior, spill-over effects from local public expenditure, tax competition across governments (see Wilson, 1999, for a recent survey), or as being the result of 'yardstick competition' among politicians (see below). Worse than that, these theories may have very different, indeed contrasting, normative implications, so that the observer is left without solid grounds to judge this evidence.

As usual in empirical analysis, these problems arise from one (or both) of the following reasons. Either alternative theories may be "observationally equivalent", or the available data set may not be rich enough to allow to discriminate among their different predictions. Consequently, solving these problems requires either to carefully re-examine the implications of the theories to be tested, or to attempt to build a better data set.

In this paper we follow both strategies, looking at potential evidence of yardstick competition across local governments in Italy. We focus on this theory for a number of reasons. First, as forcefully argued by Salmon (1987), yardstick competition theory may provide new and powerful insights on the structure of government in democratic countries. If it were really the case that citizens made comparative performance evaluation across local governments in order to understand the quality or the competence of their politicians, a new approach to key issues of fiscal federalism would follow naturally. For example, if this evaluation is beneficial to citizens, one would then want to organize, say, the allocation of functions and resources to local governments so as to maximize this behavior, possibly in contrast with the suggestions of traditional theory (as summarised, for example, by Oates, 1972, or Wildasin, 1986). A second motivation is simply that the interest in these issues is not purely academic. Indeed, several European countries – where one would expect other forms of competition across governments, such as the celebrated Tiebout's (1956) "voting by feet", to be less relevant than, say, in the United States – are actually involved in deeply reforming their local government structure.[1] Knowing if yardstick competition is a true phenomenon may then help them in designing a better institutional framework.

We begin by carefully reviewing yardstick competition theory, using the framework of Besley and Case (1995a) as a benchmark, and extending it in a number of directions. We point out that, differently from the interpretation often given in the existing literature, yardstick competition in tax setting does not necessarily yield mimicking behavior among neighboring governments. Moreover, mimicking behavior could rather be consistent with alternative theories, such as for instance tax competition. On the other hand, yardstick competition theory does yield two clear predictions. First, only incumbent governments that have electoral concerns should interact strategically with their neighbors – for example, 'lame ducks' should behave differently. Second, electoral performances should be consistently affected by tax setting both in own and in neighboring jurisdictions. This implies that in order to discriminate between yardstick competition and other competing theories, tax setting and electoral constraints, features, and outcomes should be considered at once in the empirical analysis.

To test the theory, we build a comprehensive data set on local property tax variables from a sample of Lombardian cities, that includes detailed information on local electoral histories during the 1990s. We

[1] In Italy for example the idea that decentralization and the ensuing competition across local governments may enhance efficiency in the production of local services has been determinant in shaping the 1997 tax reform and the subsequent Constitutional reform (February 2001).

argue that this setting provides a text-book case for testing yardstick competition phenomena. For the reasons indicated above, we specify and estimate both a tax setting equation and a popularity equation that allow for spatial dependence, fully exploiting both the institutional local government structure and the techniques developed within the spatial econometrics literature (Anselin, 1988; Anselin and Florax, 1995; Anselin *et al.*, 1996). As argued in section 4 below, careful spatial econometric testing, modelling and estimation are required in order to be able to identify the 'true' process generating the observed spatial pattern in the data.

Our results largely support the thesis of strategic interaction in tax setting among neighboring jurisdictions. In particular, the results of the estimation of a tax setting equation suggest that the component of the tax rate that is not 'explained' by local determinants is correlated among neighbors. While this result could simply be taken as evidence of the existence of spatially correlated 'shocks' to tax setting that have no behavioral significance, it turns out that such correlation only exists for those jurisdictions where the mayor can run for re-election and the electoral outcome is uncertain. This correlation is instead absent for those jurisdictions where either the mayor cannot run for re-election because he faces a binding term limit, or where the mayor is backed by such a large majority to make the electoral threat less poignant. These results are clearly in agreement with yardstick competition theory. Furthermore, while both mayors that can run for re-election in uncertain contests and those who have to step down because of a term limit tend to keep tax rates low in election years, thus suggesting the existence of electoral cycles (that for 'lame ducks' are probably motivated by party discipline), mayors that run for re-election and are supported by very large majorities do not tend to reduce tax rates when elections approach, a result again compatible with yardstick competition theory. On the other hand, the results of the estimation of a popularity equation show only weak evidence of a negative impact of own taxes and of a positive impact of neighboring jurisdictions' taxes on incumbents' chances of re-election. This is possibly due to the small size of the sample and the inevitable difficulties of taking into account further aspects of local policy making – such as the endogeneity of the tax variables and the simultaneous setting of multiple policy instruments.

Taken as a whole, we can conclude that our results offer some support for the existence of yardstick competition phenomena in the Italian context. Going back from empirics to theory, they also point out to some limitations of the existing approach and the need to integrate it with other observed phenomena, a point to which we will come back in the concluding section.

The rest of the paper is organized as follows. Section 2 discusses yardstick competition theory in greater detail and identifies some key predictions of the theory. Section 3 presents the data set, arguing for its validity as a natural laboratory for testing the theory. Section 4 tackles the problem of the empirical specification required to test the theoretical predictions. Section 5 presents and discusses our estimation results. Section 6 concludes by summarizing our findings, stressing the caveats of the present analysis and suggesting avenues for further research.

2. Testing Yardstick Competition theory

Strategic interaction in fiscal variables across spatially close jurisdictions may occur for a variety of reasons. Yardstick competition theory points towards one such potential link. The main idea is that neighboring jurisdictions' choices may provide voters with a positive informational externality on the quality of their own government. By observing fiscal choices in neighboring jurisdictions, citizens should be able to evaluate more precisely if the choices of their own government are appropriate, or if they include some waste – say, political rents – which citizens dislike. Hence, voters should condition their electoral decision on the observation of the relative setting of the fiscal variables in their own and in neighboring jurisdictions.

However, this simple story has several caveats. First, citizens must be able to observe fiscal choices in other jurisdictions beside their own. Second, the economic and institutional conditions in these jurisdictions must be sufficiently close to make relative comparison meaningful. Third, it has to be true that electoral behavior is the main tool citizens have at their disposal to punish bad incumbents. For example, if mobility costs are sufficiently low, instead of voting them down, citizens could escape incompetent governments by simply emigrating or moving away taxable assets. Fourth, one needs to

consider the strategic reaction of politicians to voters' behavior. Politicians, as rational agents, are aware of the kind of calculations citizens make. Hence, they might change their fiscal choices to influence voters' behavior, and one has to establish how, in order to come up with testable hypotheses linking fiscal variables to electoral results.

Beginning with the latter point, most of the literature seems to have taken for granted that, in the presence of comparative performance evaluation, bad or incompetent governments may be forced not to set taxes that put them out of line with other governments, in order to avoid being unseated. Hence, 'copycatting' in local fiscal choices will result. However, this is only one possibility. Another possibility is that by improving voters' information set, yardstick competition increases the cost for bad or incompetent governments to imitate good or competent governments, so inducing no or little mimicking behavior among different types of government.

2.1 A simple model

To bring down this point formally and discuss other implications of yardstick competition, it is useful to go back to theory. To this aim, consider the following version of Besley and Case (1995a) original model.[2] The economy has three agents: an incumbent politician, an opposing politician and a voter. The economy lasts two periods[3] and at the end of the first period an election takes place. In both periods, the incumbent politician chooses taxes and public good supply. Governments come of two types: they can be either 'good' or 'bad.' Good politicians only want to provide the public good at the lowest possible cost; bad governments like instead to tax citizens more heavily in order to accumulate rents.[4] Governments know their type, while citizens have only some a-priori on the type of governments. More specifically, we assume that:

- the citizen expects both the incumbent and the opposing government to be good with probability θ , $\theta \in (0,1)$;
- the production function of the public good is subjected to random shocks. Producing public good level g^* costs t^* if the shock is positive and $t^*+\Delta$, $\Delta>0$, if the shock is negative;
- negative shocks occur with probability $q<1$;
- all agents in the economy discount future at rate $\delta<1$.

The sequence of events is as follows. In period 1, nature moves first and chooses both the type of the incumbent and the realization of the shock. The incumbent government observes those moves of nature, and then decides the tax rate and public good supply in the first period. The citizen does not observe the moves of nature, but observes the fiscal choices of the government and knows the stochastic structure of the economy. He thus uses his observations to revise his beliefs on the type of the incumbent.

In what follows, we will indicate with $\mu(\theta, t, g)$ the posterior beliefs of the citizen. An election then takes place, and the citizen votes for the politician who expects to be good with higher probability at the time of voting. When indifferent, he votes for the incumbent[5]. That is, the incumbent is re-elected if $\mu(\theta, t, g) \geq \theta$. Whichever government is in charge in period 2, it chooses again a level of the tax rate and of public good supply. The world ends here.

To solve this dynamic incomplete information game between voter and politicians, we look for Bayesian perfect Nash equilibria of the game: that is, situations where the strategies chosen by each

[2] There are several other examples of signaling models of policy (for a survey see Persson and Tabellini, 2000, chapter 4) which could have been used to make the point, but the model of Besley and Case (1995a) has the advantage of being very simple and well known in the literature.

[3] Considering two periods only is of course just a analytic simplification. As a matter of interpretation, the last period is simply the period where incumbents do not care any longer about voters' reaction to their tax setting behavior, maybe because elections are very far off in the future or maybe because incumbents cannot run again, perhaps as a result of binding term limits. As will be clear, both interpretations are consistent with our data set.

[4] An alternative interpretation is that bad governments are simply incompetent; they have to exert an higher level of effort than good governments in order to produce the public good at the minimum cost. Nothing would change under this alternative interpretation.

[5] This assumption effectively rules out mixed strategy equilibria. See Bordignon and Minelli, (2001) and below.

agent are optimal given the strategies played by the other agents and the beliefs of the voter, and where the posterior beliefs of voters, whenever it is possible, are derived by applying Bayes' rule. In turn, the strategies of the incumbent are a choice of the fiscal variables for both the first and the second period; the strategies for the opponent are a choice of the fiscal variables for the second period, and a strategy for the voter is simply his electoral behavior, as specified above. For simplicity, we only focus on pure strategies equilibria of this game.

Again for simplicity, suppose that governments of both types have to supply a fixed amount of public good, g^*, in both periods[6]. Also suppose that good governments do not play strategically; they just do what is best for the citizen in the two possible cases; i.e. they impose a tax equal to t^* if the shock is positive and $t^*+\Delta$ if the shock is negative.[7]

Consider next the bad government and suppose that there is a maximal tax rate he can impose on citizens (an upper bound on politically sustainable expropriation) and let us write, for notational convenience, this upper tax rate as $t^*+k\Delta$, where $k>1$. The maximal rent the bad type can earn in each period is either $(k-1)\Delta$ or $k\Delta$ depending on the realization of the shock. Which tax choices does the bad politician make? In the second period, the answer is simple. As the world ends there, if he is in charge in that period, he chooses the highest possible tax rate, $t^*+k\Delta$, as this tax rate maximizes his utility. In the first period, however, he faces a trade off. He could still choose $t^*+k\Delta$, earning immediately the maximal rent, but then the citizen would immediately understand that he is of the bad type and throw him out of office at the ensuing election. Alternatively, he could try to mimic the good type and then be reelected with some positive probability. Let us suppose that the out-of-equilibrium beliefs of the voter are such that she assigns zero probability to a government being of a good type upon observing a tax rate different from one of the two rates a good government could possibly choose; i.e. $\mu(g^*, \theta, t)=0$ for $t\neq t^*+\Delta$, and $t\neq t^*$. Then it is clear that if the shock happens to be negative, the bad type would prefer to choose $t^*+k\Delta$ in the first period, even if that implies losing the election. This is so because the maximum tax rate the bad type can impose without being discovered in period 1 is $t^*+\Delta$; but if the shock is negative, this means that the bad type earns zero rents in period 1. And even under the optimistic beliefs that he would be elected for sure if he followed that strategy, he would prefer to separate immediately, since waiting is costly: $(k-1)\Delta > \delta(k-1)\Delta$. What if the shock is positive? Repeating the argument just given, the bad type would never play t^* , as his first period rents would again be zero. But could he play $t^*+\Delta$?

Proposition 1 *Under the assumptions stated above, there exists a unique perfect Bayesian equilibrium in pure strategies. In this equilibrium, if the shock is positive and* $\delta\geq\delta^\circ\equiv (k-1)/k$, $q\geq 1/2$, *the bad type first period choice is* $t^*+\Delta$. *The citizen's posterior beliefs are such that, upon observing* $t^*+\Delta$, *the incumbent is re-elected for sure.*

Proof. " By Bayes' rule, at the proposed governments strategies the citizen posterior beliefs upon observing $t^*+\Delta$ can be computed as $\mu(\theta, t^*+\Delta)= q\theta / (q\theta+ (1-q)(1-\theta))$. It follows $\mu(\theta, t^*+\Delta)\geq\theta$ if $q\geq 1/2$. Assuming this condition is satisfied, the bad type would play $t= t^*+\Delta$ if this gives him an higher expected utility than playing his favorite strategy in the first period, i.e. if $\Delta+\delta k\Delta\geq k\Delta$, which holds true if $\delta\geq(k-1)/k$. Thus, if $\delta\geq\delta^\circ\equiv (k-1)/k$, $q\geq 1/2$, and the shock is positive, in the first period good type plays t^* and the bad type $t^*+\Delta$, and they are both reelected. If the shock is negative, on the other hand,

[6] As long as the quality of the public good supplied is observable by citizens, this assumption is of no consequence and only simplifies the analysis. Note however that if we added public good quality to the model and we further assumed that this variable could be adjusted at will by the government without being observed by citizens, then only fully pooling equilibria in the tax rates would result as consequence of yardstick competition. The tax rates, in other words, would become totally uninformative on the type of government (see Bordignon and Minelli, 2001a). For the argument developed below to make sense, it is then necessary to assume at least that the room a bad government has in changing quality is limited, so that he would prefer to use taxes to accumulate rents or save effort.

[7] This assumption is without loss of generality in this context, as the good type would not have any way, under the (reasonable) restriction on citizens out-of equilbrium beliefs we impose below, to signal itself at a lower cost than the bad type.

in the first period good type plays $t^*+\Delta$ and the bad type, by dominance, $t^*+k\Delta$. The former is reelected and the latter is not. "

The intuition behind this proposition is straightforward. If the negative shock occurs with a high enough probability ($q \geq 1/2$), the citizen, although aware of the strategies played by the bad type, upon observing $t^*+\Delta$ concludes that with higher probability this comes from a good type facing a negative shock, and re-elects him. Furthermore, if the discount rate is large ($\delta \geq \delta°$), so that future matters enough, the bad type finds it convenient to mimic the good type in the presence of a positive shock, and be re-elected in the second period. Note that the larger is the maximal tax the bad type can impose in the first period (i.e. the larger is k), the larger must be the discount rate for the bad type to pool. This makes sense; as k increases, the cost of waiting increases and so must δ to support a 'pooling equilibrium,' that is, an equilibrium where in the first period the bad type of government does not choose his favorite strategy, but the one which would have been chosen by a good government in the presence of a negative shock.

To introduce comparative performance evaluation, suppose now we double our simple economy, introducing another jurisdiction which has exactly the same characteristics as the economy we just studied. That is, in both economies, governments can be of the two types just described, citizens hold the same a priori beliefs on incumbents and opponents, good incumbents do not play strategically and so on. Symmetry, of course, is only useful in simplifying the analysis; the basic insights of the model would go through even in asymmetric economies. A potential advantage for the citizen of having two economies instead of one is that she can now try to learn something about his incumbent in the first period, by looking at what is happening in the other jurisdiction. That is, the voter's posterior beliefs in region i may now be written as a function of the fiscal choices of both jurisdictions: $\mu_i(\theta, t_i, t_j)$, where the suffix indexes the jurisdiction, i=1,2, and t refers to first period choices.[8]

Of course, for comparative performance evaluation to be meaningful, it is necessary that the two economies be somehow related. Besley and Case (1995a) assume perfect correlation among regional technological shocks. Clearly this is implausible for most empirical applications of the model: if neighboring jurisdictions were actually hit by perfectly correlated (common) shocks, then it would be extremely hard to disentangle the correlation in tax rates due to common shocks from the one due to tax mimicking.

We take instead a more general approach. Let σ be the parameter which measures the degree of correlation between the two economies. Denoting by Prob(X,Y) the joint probability that region i is hit by a shock X, while region j by shock Y, where X and Y indicate the nature of the shock, we have:

$$(1) \quad \text{Prob}(N,N) = \sigma q \; ; \quad \text{Prob}(P,N) = \text{Prob}(N,P) = (1-\sigma)q \; ; \quad \text{Prob}(P,P) = 1-q(2-\sigma) \; ;$$

where N and P stand, respectively, for negative and positive shock. For $\sigma=q$ the two technological shocks are independent, for $\sigma=1$ they are perfectly correlated, while for $q<\sigma<1$ they are positively but imperfectly correlated.[9] We assume instead the draws by nature of the type of government to be independently made in the two regions.[10]

Local governments are supposed to make their tax choices simultaneously in the first period and it is therefore important to specify exactly what they know at the time of making these choices. We assume that each incumbent only observes the realization of the shock in his own jurisdiction, which seems reasonable. Concerning the type of the neighboring government, Besley and Case (1995a) assume that

[8] We suppress *g* from the argument of the revised beliefs' function because public good supply is assumed to be fixed.

[9] Negative correlation is also a theoretical possibility, but for the problem at hand positive correlation is clearly more plausible.

[10] One could question this assumption on the grounds that political affiliation of local governments is likely to be spatially correlated and that political affiliation may also be correlated to governments' quality. Interestingly, in our data set there is no evidence of such spatial correlation and party affiliation turns out to be insignificant in our regressions. See section 5.

each incumbent politician observes the choice by nature of the other local government type. This assumption does not seem plausible, however. Politicians probably know more than the layman about the quality of neighboring politicians, but it is unlikely that such information extends to perfect knowledge of the other incumbent type. In what follows, we will then make the opposite assumption that each incumbent only knows the distribution over government types of the other jurisdiction. As this assumption is also extreme, we will comment briefly below of the consequences of relaxing it and retaining the original Besley and Case (1995a)'s formalization.

The central question to address is what effect comparative performance evaluation has on the equilibria of the game and therefore on the empirical predictions of the model. To see this, note first, by using the same dominance argument we used before, that if the shock is negative, bad governments will again prefer to separate in the first period by choosing the maximum tax rate and being defeated at the election. By the same token, if the shock is positive, they will never play t*, as their first period rents would again be zero. Would then a bad politician play $t^*+\Delta$?

To answer this question, note that at an equilibrium where all agents expect the bad types to play $t^*+\Delta$ in the first period when the shock is positive, the expected probability of being reelected for a bad type in region i if he follows this strategy, let us call it $G(t^*+\Delta)$, is:

$$(2)\ G(t^*+\Delta) = \text{Prob}(\varepsilon_j = N \mid \varepsilon_i = P)\ (\theta\, R(t_i=t^*+\Delta,\ t_j = t^*+\Delta) + (1-\theta)\, R(t_i=t^*+\Delta,\ t_j = t^*+k\Delta)) + \\ + \text{Prob}(\varepsilon_j = P \mid \varepsilon_i = P)\ (\theta\, R(t_i=t^*+\Delta,\ t_j = t^*) + (1-\theta)\, R(t_i=t^*+\Delta,\ t_j = t\ t^*+\Delta)).$$

$$= \{(1-\sigma)q\ (\theta R(t_i=t^*+\Delta,\ t_j = t^*+\Delta)) + (1-\theta)R(t_i=t^*+\Delta,\ t_j = t^*+k\Delta)) + \\ (1-q(2-\sigma))(\theta R(t_i=t^*+\Delta,\ t_j = t^*) + (1-\theta)R(t_i=t^*+\Delta,\ t_j = t^*+\Delta))\}/(1-q)$$

In (2) the expectation is taken both with respect to the conditional probability about the technological shock the neighboring government is facing (which the incumbent in region *i* can compute having observed the realization of his own shock), and with respect the realization of the type of this government. $R(t_i=t^*+\Delta,\ t_j = t)$ is the probability of being reelected by the bad type in region i if he plays $t^*+\Delta$ and government j plays t, $t= t^*,\ t^*+\Delta,\ t^*+k\Delta$. Clearly, by our previous assumption, $R(t_i=t^*+\Delta,\ t_j = t)=1$ if $\mu(\theta,t_i=t^*+\Delta,\ t_j = t)\geq \theta$ and $R(t_i=t^*+\Delta,\ t_j=t)=0$ otherwise. By applying Bayes' rule at proposed strategies for both types of government, it is an easy matter to compute those ex-post beliefs[11] as follows:

(3)

$$\mu(\theta,\ t_i=t^*+\Delta,\ t_j = t^*+\Delta) = \{\sigma q\theta^2 + \theta(1-\theta)(1-\sigma)q)\}/\ \{(\sigma q\theta^2 + 2\theta(1-\theta)(1-\sigma)q + (1-\theta)^2(1-q(2-\sigma))\}.$$

$$\mu(\theta,\ t_i=t^*+\Delta,\ t_j = t^*+k\Delta) = \sigma\theta\ /\ \{\sigma\theta + (1-\theta)(1-\sigma)\}.$$

$$\mu(\theta,\ t_i=t^*+\Delta,\ t_j = t^*) = s\theta\ /\ \{s\theta + (1-\theta)(1-s)\}. \qquad \text{where } s= (1-\sigma)q/(1-q).$$

It follows:

(4.1) $R(t_i=t^*+\Delta,\ t_j = t^*+\Delta) = 1$ if $\theta + 2(1-2\theta)(1-\sigma)q \geq (1-q)$, $R(t_i=t^*+\Delta,\ t_j =t^*+\Delta) = 0$ otherwise;

(4.2) $R(t_i=t^*+\Delta,\ t_j = t^*+k\Delta) =1$ if $\sigma\geq 1/2$, $R(t_i=t^*+\Delta,\ t_j =t^*+k\Delta) =0$ otherwise;

(4.3) $R(t_i=t^*+\Delta,\ t_j = t^*) =1$ if $(1-\sigma) \geq (1-q)/2q$, $R\ (t_i=t^*+\Delta,\ t_j =t^*) =0$ otherwise.

[11] In deriving (3) we use the fact that at the proposed strategies for the two types of government, if the voter in region i observes that in region j the tax chosen is t* ($t^*+k\Delta$) he can immediately infer that in region j the government is good (bad) and the shock has been positive (negative).

Using condition (4), we can answer the question raised above. We begin by establishing the following:

Proposition 2. *"Suppose $q<1/2$, $\sigma>1/2$ and $k<k^*$. Then for $\theta\in[\theta^*, 1)$ and $\delta\in[\delta^*, 1)$ there exists a unique perfect Bayesian equilibrium in pure strategies where bad type's first period choices in both economies upon observing a positive shock are $t^*+\Delta$."*

Proof. "Invoking (4), under the conditions stated in the proposition, at the proposed equilibrium strategies for the two types, voter's revised beliefs in region i, i=1,2, j=1,2 $i\neq j$, are such that $R(t_i=t^*+\Delta, t_j = t^*+k\Delta) =1$ and $R(t_i=t^*+\Delta, t_j = t^*) =0$. $R(t_i=t^*+\Delta, t_j = t^*+k\Delta) =1$ if $\theta \geq\theta^* \equiv ((1-q(3-2\sigma)) /((1-4(1-\sigma)q))<1$ as $\sigma>1/2$. Invoking (2), $G(t^*+\Delta)= \{(1-\theta)(1-q) + \theta q(1-\sigma)\}/(1-q)$. The expected utility of the bad type when playing his equilibrium strategy is then $\Delta+\delta k\Delta\{(1-\theta)(1-q) + \theta q(1-\sigma)\}/(1-q)$. This dominates his best deviation, playing $t^*+k\Delta$ and not be reelected in the second period, if $\delta \geq\delta^* = (k-1)(1-q) /k\{(1-\theta)(1-q) + \theta q(1-\sigma)\}$. In turn $\delta^*<1$ if $k <k^*\equiv(1-q)/ \theta(1-q(2-\sigma))$."

A comparison between propositions 1 and 2 makes it clear that it is in principle possible to find values of the parameters that support a pooling equilibrium under yardstick competition when this would not be possible otherwise (as $q<1/2$). Intuitively, if the correlation between the two economies is large enough and the reputation of the incumbent is also large enough, observing $t^*+\Delta$ or $t^*+k\Delta$ in region j reassures the voter in region i that effectively a bad shock occurred in her region and that therefore the incumbent playing $t^*+\Delta$ in his region is a good government. The condition on the discount rate is of course more restrictive than that in proposition 1 ($\delta^* >\delta°$), because the bad type is now unsure about the type of the incumbent in the other region and so the expected probability of being reelected is smaller when he plays $t^*+\Delta$. However, this is a result of our assumption about the information held by a government about the other incumbent. As is easy to check, under the same conditions on parameters stated in proposition 2, the probability of a bad type to be reelected when playing $t^*+\Delta$ is again one if he knows for certain to face a bad incumbent in the other region, so that in this case the condition on the discount rate would be the same as in proposition 1.[12] Vindicating the empirical literature mentioned above, it is also easy to check that under the conditions stated in proposition 2 the ex ante probability[13] of finding the two regions selecting similar tax rates (in the first period) would be higher under yardstick competition than without it, so that a high correlation among the tax rates of neighboring regions (in the first period) could be interpreted as a signal of the presence of comparative performance evaluation among governments by voters.
This is not the only possibility, however, as the next proposition illustrates.

Proposition 3 *Suppose $q\geq1/2$ and $(1-q) >\theta$. Then, there exist $q\leq\sigma_1 <\sigma_2<1$ such that the bad type's first period equilibrium choices in both economies upon observing a positive shock are $t^*+\Delta$ if the following conditions are satisfied: (i) $\sigma>\sigma_2$, $\delta\geq\delta^{**}$ and $k<k^{**}$; (ii) $\sigma_2 \geq\sigma>\sigma_1$, $\delta\geq\delta^*$ and $k<k^*$; (iii) $\sigma_1\geq\sigma$, $\delta\geq\delta°$, where $\delta°<\delta^*<\delta^{**}<1$ and $1<k^{**}<k^*$.*

Proof. "At the proposed strategies for the two types, voter's posterior beliefs are such that
(1)$R(t_i=t^*+\Delta, t_j = t^*+k\Delta) =1$ $\forall\sigma$, as $\sigma\geq q$ and $q\geq1/2$ by assumption;

[12] On the other hand, if the bad incumbent in a region knew that a good government is in charge in the other region the conditions to support a pooling equilibrium would become more restrictive than those stated in proposition 2 (the discount factor should be even higher) and a pooling equilibrium could not be supported at all for $\sigma=1$.

[13] Under the conditions stated in proposition 2, the ex-ante probability of finding in the first period a difference equal to the maximal possible difference in the tax rates selected by the two regions ($k\Delta$) is equal to $2\theta(1-\theta)q$ if voters do not make comparative evaluation and it is equal to $2\theta(1-\theta)q (1-\sigma)$ if they do make comparative evaluation. Furthermore, the ex-ante probability of finding two identical tax rates in the two regions is equal to $\theta^2(1-2q+2q\sigma) + (1-\theta)^2$ if voters do not make comparative evaluation, and it is equal to $\theta^2(1-2q+2q\sigma) + (1-\theta)^2(1-2q+2q\sigma) +2\theta(1-\theta)q (1-\sigma)$ if they do make comparative evaluation. The latter is larger than the former if $\theta>1/3$, which holds certainly true since $\theta^*>1/3$.

(2) $R(t_i=t^*+\Delta, t_j = t^*) =1$ if $\sigma \leq \sigma_1 \equiv (3q-1)/2q$.
(3) $R(t_i=t^*+\Delta, t_j = t^*+\Delta) = 1$ if $\sigma \leq \sigma_2 \equiv \{(3q-1) + \theta (1-4q)\} /2q(1-2\theta)$. Note that $q\geq 1/2$ and $(1-q) >\theta$ implies $\theta<1/2$ so that $\sigma_2 >0$. Furthermore, $q\geq 1/2$ also implies $\sigma_1 \leq \sigma_2$.
Using (1)-(2)-(3), it is clear that in case (iii), the expected probability to be elected for the bad type when playing $t^*+\Delta$ in the first period upon observing a positive shock is just one, so that the condition on δ for the existence of a pooling equilibrium coincides with those in proposition 1, $\delta\geq\delta°$; in case (ii), this probability is $\{(1-\theta)(1-q) + \theta q(1-\sigma)\}/(1-q)$, so that the conditions on δ and k for the existence of a pooling equilibrium coincides with those of proposition 2, $\delta\geq\delta^*$ and $k<k^*$. In case (i), finally, $G(t^*+\Delta)= \{(1-\theta)q(1-\sigma)\}/(1-q)$, giving an expected utility to the bad type when playing $t^*+\Delta$ in the first period of $\Delta+\delta k\Delta\{(1-\theta)q(1-\sigma)\}/(1-q)$. This dominates playing $t^*+k\Delta$ in the first period if $\delta \geq\delta^{**} = (k-1)(1-q) /k\{(1-\theta)q(1-\sigma)\}$. In turn $\delta^{**}<1$ if $k <k^{**}\equiv(1-q)/ (1-q(2-\sigma(1-\theta) - \theta)$."

Clearly, in this second example, supporting a pooling equilibrium becomes unambiguously more difficult under yardstick competition than without it, in the sense that stricter restrictions on the parameters must be imposed to get pooling behavior. Furthermore, these restrictions increase with the degree of correlation between the two economies. When this correlation is very low, $\sigma \leq\sigma_1$, a pooling equilibrium can be supported under the same conditions as in the case where there is no comparative performance evaluation. As σ increases, δ must also increase to support a pooling equilibrium, and this equilibrium simply disappears when the two economies become perfectly correlated. Intuitively, the larger the degree of correlation, the more the citizen learns by observing the fiscal choices in the other jurisdiction and the more difficult is for the bad type to escape detection when cheating in the first period[14]. Again, none of these qualitative results depends on our assumption on the information held by a politician on the type of the politician in the other jurisdiction (although the specific values of the parameters supporting the pooling equilibrium do). As is easy to check, even if a bad incumbent in a region knew for sure that another bad type is in power in the neighboring jurisdiction, under the conditions stated in proposition 3, a pooling equilibrium could be sustained only at the cost of higher values of the discount parameter as σ becomes large. Empirically, in this second example, one should *not* expect to find positive correlation in tax rates among neighboring jurisdictions as a result of yardstick competition.

2.2 Empirical predictions of Yardstick Competition theory

Propositions 2 and 3 clearly do not exhaust all the possibilities of analysis of the model,[15] but they are enough to establish the general point that depending on the parameters, yardstick competition may induce either more pooling or more separating behavior among different types of government (in the first period). In principle, if one knew these parameters' values, it could be still possible to derive form

[14] Interestingly, note that there is not a monotone relationship between the degree of correlation and the kind of equilibria which may result. A higher level of correlation makes it easier to support a pooling equilibrium when this would not be sustainable without comparative performance evaluation competition in proposition 2, but it makes it more difficult in proposition 3. In a sense this points to the intrinsic limitations of the electoral mechanism as a way of disciplining policy makers. As is well known from game theory, even a tiny bit of correlation between agents' types is enough to reach the first best for the principal (the voter in this case) as long as unlimited punishments can be imposed on agents. But in the case of the electoral system, the only possible punishment (reward) on politicians is not to re-elect (re-elect) them.

[15] In particular, to save space we refrain from asking in the text if comparative performance evaluation makes the voter better or worse off. It can be shown that if the rents the bad types can raise are large enough (k>2), yardstick competition benefits the voter when a pooling equilibrium can still be supported even under comparative evaluation and makes him worse off in the opposite case. Intuitively, when the maximal rents the bad government can raise are large, the cost of learning the type of government is too high for the voter, because he has to pay those rents sooner, and therefore a separating equilibrium reduces welfare. In a pooling equilibrium, on the other hand, since the government is discovered more easily at no extra cost, the voter is better off. Although seemingly paradoxical, these results should not come as a surprise. Indeed, they are just instances in this framework of the more general problem concerning 'career concern model' stressed by Meyer and Vickers (1997).

the theory precise implications on the tax setting behavior of neighboring jurisdictions which could then be tested. In practice, however, this is likely to be extremely difficult. Furthermore, as noted already in the Introduction, copycatting in tax rates could well be consistent with other theories. For example, if the tax is levied on a spatially mobile asset, attempts to raise the tax rate in a jurisdiction would lead to a loss of the tax base. If governments were aware of that, they would select strategically their tax rates so as to attract tax revenue. In a non-cooperative equilibrium in tax setting, one would again expect to find spatially close jurisdictions to choose similar tax rates.

It is then clear that if one wants to test yardstick competition theory, one needs to go beyond the standard analysis of tax setting behavior. In fact, the main focus of yardstick competition theory is not on tax setting as such, but on tax setting as connected with the existing electoral constraints. Looking at the relationship between the two, one can search for predictions which are specific to yardstick competition theory. On these grounds, note first that a clear and specific prediction of the theory is in term of a quite different behavior between incumbent politicians who run for a second period in office and incumbents who do not – as already argued by Case (1993) and Besley and Case (1995a; 1995b). Since the only reason to mimic a good government in the first period is to be elected in the second, bad governments should not be too worried about competing with their neighbors if there is no such second period, for instance because of the existence of a binding term limit. Moreover, the theory also suggests that only incumbent governments that face 'uncertain' electoral outcomes should interact strategically with their neighbors: if an incumbent government is pretty confident of re-election in spite of his tax setting behavior – for example because he knows he is supported on other grounds by a large majority of the population, or is not challenged by any credible contender – we should again not expect to find his fiscal choices to be affected by his neighbors.'[16]

To put it differently: yardstick competition may or may not induce positive spatial auto-correlation in tax rates among neighboring jurisdictions, as we saw above. However, if we observed a correlation in tax rates *only* among neighboring jurisdictions ruled by politicians in their first term of office that face a serious chance of being unseated, this could be interpreted as prima facie evidence corroborating yardstick competition theory, in the sense of being at least consistent with some possible outcomes of this theory, and not with other theories. Indeed, if correlation in tax rates were only the result of tax competition or other spatially related phenomena, it would be very hard to explain why there should be such a heterogeneous behavior between 'first period' and 'last period' governments.

By the same token, if we did not observe or observed little correlation in tax rates among neighboring jurisdictions, this could not be taken as conclusive evidence of the absence of yardstick competition among local governments, as this result could also be consistent with some parameterizations of the model. In this case, however, theory makes an other precise prediction: relative tax setting should affect the reelection chances of incumbents, in the sense that politicians choosing higher tax rates than their neighbors should lose electoral support. Note again that this prediction is specific to yardstick competition theory: it is hard to think of any other story which could be consistent with this behavior. Case (1993) and Besley and Case (1995a) capture this idea by regressing the probability of U.S. state gubernatorial defeat against changes in state income tax liabilities relative to neighboring states. In what follows, given the institutional characteristics of local governments in Italy, we prefer to capture it in terms of a 'popularity function' of incumbents at the municipal level, that allows us to fully exploit the electoral information that is available in our data set.

[16] To save space we do not provide a formalization of this prediction here, but this could be done very easily. Suppose for example that population in the jurisdiction is now divided in two groups, 'active' and 'passive' voters. Passive voters only vote ideologically, while active voters only care about the tax setting behavior of the incumbent, along the lines of the model discussed in the previous section. Suppose further that while active voters always vote, passive voters' electoral turn out is a random variable, with support large enough to make the incumbent never sure to win the elections. It can then be shown that if the share of passive voters in the population is large enough and the ideological advantage of the incumbent among passive voters is also large enough, a bad incumbent would never choose to mimic a good one in the first period, as the expected benefit of doing so in terms of a higher probability of winning the election is always lower than the cost of giving up his favorite strategy in the first period. Details of this extension are available by the authors on request.

3. The data set: property taxation and politics in a sample of Italian cities

In order to explain how we capture these key predictions of the theory in our empirical analysis, it is useful to begin by presenting first the institutional scenario. We build a comprehensive data set on local property taxation in a sample of municipalities in Lombardia, a large region in the north of Italy, including detailed information on those cities' electoral histories during the 90's. As already hinted in the Introduction, we choose this setting because we believe it provides a very suitable framework for a test of yardstick competition phenomena.

The local property tax rate, named ICI (*Imposta Comunale sugli Immobili*), was introduced in 1993 together with a massive reform of the electoral system of municipalities, with the specific aim to increase the administrative power and the accountability of city governments. Since then, the decision about the property tax rate has become the most relevant financial decision an Italian mayor must make each year; on average, property tax revenue alone covers more than 50% of total tax revenues of local governments and more than 25% of total revenue. The mayor is also uniquely accountable for this decision; as we explain in more detail in the next section, since the 1993 electoral reform, Italian mayors are directly elected by citizens in a winner-takes-all ballot which provides them with at least 60% of the seats in the City Council and the upper hand on city government.[17] There is also no question about the visibility of this decision; the local property tax rate is virtually observed by everybody for the simple reason that a large majority of Italians are home-owners, and therefore have to pay this tax. To be more specific, ICI applies to both "domestic" and "business" property. The domestic property tax rate is applied to resident household owners only; the business property tax to all the rest of buildings, housing, shops or offices. In both cases, the property tax base is defined by national procedures and regulations and it is therefore uniformly determined across cities and towns. Municipalities are free to choose both tax rates separately in the range from 0,4 to 0,7 percentage points on the tax base of the property, although they usually tend to choose the same tax rate for both domestic and business taxation. However, differently from business property taxation, domestic property taxation may be accompanied by lump sum deductions chosen at will by each municipality according to specific criteria (such as the number of people living in the family, the family income and so on); it is therefore often the case that statutory domestic ICI tax rate does not coincide with the effective domestic ICI tax rate. Unfortunately, we do not have enough information on the tax bases of municipalities to be able to recover the effective domestic ICI tax rate imposed in each municipality; and moreover, even if we could do it, it is doubtful that it would be very useful for our aims, as it is hard to believe that the average citizen himself had the time or the information needed to compute these effective tax rates and to use them for comparative evaluation across local governments.[18]

For this reason, although it might have been preferable in principle to employ the domestic tax rate for our analysis (as we would have been then sure that the taxpayer is also a voter in the same jurisdiction), in what follows we use instead the business ICI tax rate, for which comparison of tax rate levels across municipalities is both easy (given the small size of Italian municipalities) and meaningful (as the tax base is commonly determined by the Central government). Note however, that given the economic structure of Italian towns, characterized by an overwhelming presence of small business in the economy, this should not be a major problem for our analysis as in most cases payers of the business ICI in a municipality are to a large extent also likely to be resident in the same municipality. Note further that in our sample revenues from business ICI are on average more than 70% of total ICI revenues.

As for the sample of municipalities we use here, it comprises 143 adjacent Comuni (municipalities) in the Province of Milan. The characteristics of the Comuni in the sample are described in table A.1 in appendix 1, from which it emerges that those municipalities differ considerably in terms of size,

[17] A mayor can change at will the composition of her government; on the contrary, a mayor cannot be replaced by its majority without undergoing new elections. The electoral rules, however, are slightly different for small and large municipalities; see below for more details.

[18] Indeed, our estimations show that there is no evidence of correlation for the statutory domestic ICI tax rates across neighboring jurisdictions, suggesting that both governments and citizens do not consider the relative comparison of these tax rates as meaningful. See also note 31 below.

population, demographic and socio-economic characteristics, even after excluding jurisdictions with less than 4000 residents. Beyond ICI revenue, the rest of local current public spending is funded by other minor own sources of revenue (charges and fees) and by grants from central government. The latter are unaffected by the property tax rates chosen by local authorities, and we will therefore consider them as exogenous variables in the ensuing analysis. The data source of all these variables, except for grants and tax rates, is ISTAT (National Institute of Statistics) and it refers to the calendar year 1991 (National Census of Government). Property tax rates have been provided by the Association of Italian City Councils (ANCI); data on grants come from the Budget Office of the Regional Government and refer to the budgetary year 1996. To avoid obvious endogeneity problems, we refrain from including in our data set budgetary information on current or capital expenditure by our municipalities, although they were available from the last source for the same year. On theoretical terms, the lack of any information on local public expenditure constitutes however a problem; we will come back to it again in the conclusions.

Table A.1 also includes information on a number of variables describing the political complexion of the cities in the sample. In particular, we collected information on the political affiliation of local governments (the coalition of parties supporting a mayor) and their electoral performances in the last five years, based on Home Ministry of the Italian Government official data. Political affiliation varies widely in the sample. Looking at the current political situation, about 55% of local authorities are controlled by centre-left governments, about 36% from centre/conservative governments, and the remaining municipalities are controlled by right-wing or Northern League parties. A mayor is appointed for 4 years (5 since the electoral reform of 1999), although it sometimes happens – even after the 1993 reform which has largely increased the stability of local governments – that a city government collapses before the terminal period. Italian mayors are subjected to a binding term limit: they cannot run for office for more than two terms in a row.

Finally, 97 of the municipalities in the sample held elections in the years 1998-2000, and for these municipalities we collected complete information on election results (vote shares earned by political parties and coalitions supporting a candidate in two consecutive elections) as well as the crucial information on the eligibility of the mayor in power to run again at the next election. As Table A.1 shows, about 75% of the mayors in power did run again in the period 1996-2000.

4. Empirical specification and the choice of variables.

In developing our empirical analysis we proceed in three steps. First, we look for evidence of local strategic interaction in tax setting. Assessing this on empirical grounds is a tricky issue. Spatial auto-correlation in tax rates among neighboring jurisdictions may occur for a variety of reasons, and it is important to be able to disentangle genuine strategic interdependence from the effect of auto-correlated shocks with no behavioral meaning. The main problem lies in that the two processes have the unfortunate property of 'mimicking' each other (Revelli, 2001b). Second, we allow for heterogeneous behavior between mayors in their first or second term of office. In particular, we use the information available in our data set to verify if the mayors who should not be concerned with election outcomes – either because they cannot run for re-election due to binding term limits, or because they are backed by extremely large majorities – do actually behave differently in their tax setting decisions from the remaining mayors.[19] Finally, to test for the other key prediction of yardstick competition theory, we explore whether local incumbents' electoral fortunes are actually interdependent, that is, whether the election outcome in a jurisdiction depends both on own and on neighbors' tax rates. To this aim, we formulate and estimate a tax setting equation and a popularity equation.

4.1 Tax setting equation

[19] Case (1993) and Besley and Case (1995a) allow U.S. state governors ineligible to run for re-election to respond differently to changes in neighbors' tax increases than those who may wish to run again. However, their empirical analysis of the tax setting equation relies on an instrumental variables principle that might not be able to properly identify the source of spatial auto-correlation, that is to discriminate interaction in the dependent variable from interaction in the unexplained component of the equation.

In the *tax determination equation*, the tax rate set by a local jurisdiction at a given time is related to a number of local variables – including the traditional 'internal' determinants of local tax setting policies (Wildasin, 1986) – as well as to neighboring jurisdictions' tax rates. In particular, the local property tax rate for the year 2000 – a (143×1) vector t of municipalities in our sample – is assumed to depend on a series of local characteristics (the matrix of explanatory variables, X), including (see Table A.1):

a) structural characteristics (area, population, urbanization rate);
b) socio-demographic characteristics (percentage of children and elderly people, rate of unemployment);
c) fiscal variables (grants from central government, disposable income per capita, tax base indicators[20]);
d) political variables (political control dummies, election year dummies and term limit dummies).

Moreover, in order to account for spatial interaction among neighboring jurisdictions, we need to include a measure of the tax rates set by close-by jurisdictions as a determinant of a given jurisdiction's tax rate:

e) neighboring governments' property tax rates.

Estimation of a tax setting equation that allows for an effect from neighbors' policies poses a number of econometric problems, mainly deriving from the fact that own and neighbors' choices are made simultaneously. Moreover, spatial dependence between neighboring governments' tax rates might simply arise from the presence of common fiscal shocks, that one could easily mistake for substantive interaction (Brueckner, 1998). Notice further that the issue of disentangling the correlation due to actual mimicking behavior from the one that is simply due to common shocks is crucial in this context, because – as shown in section 2 – the very existence of spatially correlated shocks to tax setting policies is itself a necessary condition to make comparative performance evaluation meaningful.

We will consequently proceed as follows. First, if we did not allow for the variables at point e) above, that is in the absence of any consideration of potential strategic interaction among neighboring governments, the model of local tax determination could be expressed in matrix form in a straightforward conventional way as:

$$\text{Model (C):} \qquad t = Xb + u$$

where b is a vector of parameters to be estimated and u is an error term, that is usually assumed to be independently and identically distributed across the observations. The above model can be estimated by ordinary least squares (OLS).

However, in the presence of local strategic interaction due to yardstick competition or other phenomena, model (C) would be mis-specified. Estimates of the parameters b would be biased if local governments affect each other in tax setting and we omit neighbors' taxes from the model. If strategic interaction is relevant, a model of local tax determination equation should instead be written as a "spatial lag dependence" model (Anselin, 1988):

$$\text{Model (L):} \qquad t = Xb + pWt + e$$

where p – with $|p|<1$ for spatial stationarity (Anselin, 1988) – is the parameter measuring the interaction between own taxes (t) and a measure of neighbors' taxes (Wt). In the applied spatial econometrics literature it is customary to build up a 'spatially weighted average' of neighbors' taxes in order to control for spatial interaction. This is done by premultiplying the tax vector t by a square matrix W that reflects the location – or some other 'proximity' criterion – of the jurisdictions. In particular, with N local jurisdictions, matrix W is a (N×N) matrix of spatial weights that assigns 'neighbors' to each local government: if jurisdictions i and j are neighbors according to the specified

[20] Specifically, these indicators refer to information on the number of houses in each municipalities that are not inhabited by the owner-occupants, in order to control for the effect on the different shares of residents on property tax rate setting.

criterion, the element w(i,j) in matrix W is positive, otherwise it is zero.[21]
On the other hand, a second source of mis-specification of model (C) could be due to spatial auto-correlation in the "unexplained" or "unpredicted" component of the local tax equation (Brueckner, 1998). If the error term of equation (C) is not i.i.d., but is correlated across local jurisdictions, the tax setting model should be written as a 'spatial error dependence' model (Anselin, 1988), that is as a first order spatial auto-regressive process in the error term:

$$\text{Model (E):} \quad \begin{aligned} t &= Xb + u \\ u &= rWu + e \end{aligned}$$

where r – with $|r|<1$ for spatial stationarity – is the spatial auto-correlation coefficient, and W is as defined above. Spatial auto-correlation in u could arise from either of the following sources. First, even though model (C) is correctly specified – that is, even if it does not omit relevant determinants of property tax rates – it might still be the case that local governments are hit by unpredictable and unobservable spatially auto-correlated shocks that affect their tax setting decisions. Second, model (C) could actually have been mis-specified, in the sense that a relevant determinant of the tax rate that shows itself spatial auto-correlation – say, ideological complexion of the jurisdiction, or income distribution – has been omitted from the matrix of explanatory variables X. Third, and what is particularly important in this context, spatial auto-correlation in u might emerge if comparative performance evaluation of incumbents is not made on 'raw' tax rates relative to neighbors,' but rather on "unanticipated" tax rates, that is on relative tax rates *conditional on own and neighboring jurisdictions' characteristics* (Besley and Case, 1995a). In other words, if voters take into account both their own and their neighbors' characteristics when comparing their respective tax rates, they should punish the incumbent government only for the tax rate excess that cannot be justified in terms of local preferences, needs and economic conditions. For instance, if a jurisdiction is hit by a clearly discernible idiosyncratic shock that forces it to raise taxes – say, a sudden tax base fall determined by an environmental catastrophe (a flood) – this should be no convincing argument for the taxpayers in neighboring jurisdictions that taxes should rise everywhere. Incumbents might consequently afford to imitate only that part of their neighbors' tax rises that cannot be anticipated. According to the above argument, we should expect model (L) to be rather compatible with a tax competition framework than with a yardstick competition one, with the latter yielding the sort of correlation described by model (E).
Clearly, though, we need a further identifying condition to be able to discriminate between the sheer presence of common shocks and actual "conditional" copy-catting behavior – that are observationally equivalent. In order to do so, we exploit the institutional characteristics of local government structure in Italy, in particular the existence of term limits. Furthermore, we construct a measure of incumbents' confidence of re-election in terms of the vote share earned in the previous elections. This allows us to predict that while we should observe spatial error auto-correlation irrespective of the mayors' perspectives – first or second term of office – and of their perceived degrees of popularity if the first interpretation (the existence of common shocks) is true, we should expect to find spatial error auto-correlation due to actual copy-catting behavior only for those mayors that wish to run for re-election in

[21] Yardstick competition theory hinges on the existence of economically related 'neighboring jurisdictions' whose fiscal choices are easily observed by voters. Identifying those jurisdictions on empirical grounds is not an easy matter, however. In the ensuing analysis, we follow the literature in identifying neighboring jurisdictions with spatially close jurisdictions, on the basis of the argument that information flows more easily across close-by jurisdiction and that adjacent jurisdictions are also more likely to experience similar economic shocks. Thus, the weights matrix W we use in the analysis is based on a simple geographical criterion: two municipalities are held to be neighbors if they have a common border. However we checked for the validity of this criterion by building weights matrices based on alternative criteria: similarity in per capita income levels, population size, unemployment and urbanization rates. Neither of them led to an improvement in the estimation results upon the basic geographic (border sharing) specification (results are available from the authors on request). Still, it is important to remember that testing for yardstick competition implies testing simultaneously for the particular criterion chosen for identifying neighbors. Indeed, in one of the earliest application of the theory, Case *et al*. (1991) use ethnical composition, besides spatial proximity, as an indicator of neighboring across the US states.

an uncertain contest.

4.2 Popularity equation

To further explore whether yardstick competition is really going on, the theoretical framework in section 2 suggests that we should also analyze the impact of own and neighbors' taxes on election outcomes. Since our data set contains detailed information on municipalities' election results in a number of years, we are able to estimate a *popularity equation*.

The popularity equation – equation (5) below – uses as dependent variable the local election outcome of the incumbent government – expressed as the share of the vote earned at an election held at time t+1 by the party (or coalition) that got into office at time t. We denote such variable by S(t,t+1).

Alternatively, one could adopt a discrete choice approach and use as dependent variable the probability of re-election (or defeat, as in Case, 1993, and Besley and Case, 1995a) of the incumbent (1 in case of re-election, 0 in case of defeat). However, since this would imply giving up a large amount of available information on actual vote shares earned by political parties in local government elections, we prefer to fully exploit the data set and use the continuous vote share variable on the left hand side of the equation.

The set of explanatory variables on the right hand side of equation (5) includes own and neighbors' property tax rates – denoted by t and Wt respectively – and a number of socio-economic local characteristics (matrix D). Moreover, and following some recent empirical research on vote functions,[22] we include the vote share obtained by the incumbent in the previous election – denoted by S(t,t) – to allow for an influence on time t+1 election from the previous election (persistence of shocks to popularity) and we control for the impact of national politics on local elections (NP). The latter variable is a party-specific effect that allows for a uniform impact on all local councils controlled by a given party or coalition.

$$(5) \qquad S(t,t+1) = \beta_0 + \beta_1 S(t,t) + \beta_2 t + \beta_3 Wt + \beta_4 D + \beta_5 NP + v$$

Clearly, the error term v of equation (5) could be correlated with the error term in the tax setting equation, as theory implies that local public officials might set fiscal policies strategically and opportunistically. For instance, an incumbent seeking re-election in a time of low popularity can attempt to gain votes by cutting the local tax rate. Estimation of the popularity equation must therefore control for endogeneity of own (t) and neighbors' (Wt) tax variables.[23]

When estimating the popularity equation (5), there is an important issue to be considered first. As we already hinted above, the electoral system of municipalities in Italy differs according to the size of the population. In municipalities with more than 15,000 residents, there is a two election rounds system: if a candidate gets at least 50% of the votes in the first round, he is elected mayor. If in the first election round no candidate gets at least 50% of the votes, the two most voted candidates run in a second electoral round, where further aggregations between parties (the parties supporting candidates that did not make it to the second round) usually take place. In municipalities with less than 15,000 residents, on the other hand, there is a one-round election and the mayor is the candidate that gets the highest percentage of the votes, with the coalition supporting the mayor getting two thirds of the council seats. While the latter jurisdictions do not pose particular problems, since maximization of pluralities is clearly a sensible hypothesis, things are more ambiguous in large jurisdictions due to the presence of two electoral rounds. For large municipalities, we decided to take as dependent variable in the popularity equation the percentage of votes earned in the first round by the incumbent government: we take the outcome of the first electoral round as a sincere indicator of the incumbent's popularity, while the second round percentage of the vote got by the incumbent might be driven by other unpredictable factors (such as the characteristics of the opponent). Finally, in the case where the mayor in power

[22] Reviewed in Paldam (1997).

[23] Similar comprehensive politico-economic models that explicitly recognize that election outcomes and incumbents' fiscal choices are determined simultaneously have been used in the large political business cycle and voting function literature at the national level (Paldam, 1997, for a recent review).

steps down, we consider the proportion of the vote earned in the first round of the next election by the coalition of parties that supported the winning mayor in the previous election.

5. Results

5.1 Tax setting equation

If local governments do not set taxes in isolation, but affect each other in making their fiscal choices, we should expect to find that tax rates in neighboring jurisdictions are correlated. Consequently, and in order to convince ourselves that a spatial framework is required to analyze local tax setting, we have preliminary performed a spatial auto-correlation test on the OLS residuals of a non-spatial tax setting model such as model (C). According to the Moran spatial test (Anselin, 1988), the residuals appear to be positively spatially auto-correlated, and the test is significant at the 99% level of confidence.[24]

As a result, a model of local tax determination where spatial interaction is explicitly allowed for, either model (L) or model (E), should be used instead of model (C). Unfortunately, the Moran test on the OLS residuals is unable to establish which of the two spatial models is appropriate to explain local tax setting decisions. A more precise indication of what is the most likely source of spatial dependence is given by the two robust Lagrange Multiplier (LM) tests on the OLS residuals of model (C), developed by Anselin *et al.* (1996). The LM test for spatial error dependence (that is robust to the presence of spatial lag dependence) suggests at the 95% level of confidence that the errors in the tax determination equation are spatially auto-correlated.[25] On the other hand, the LM test for spatial lag dependence does not point conclusively towards the presence of substantive dependence in the tax rates. The best way to discriminate among the two competing models, then, is to estimate them by Maximum Likelihood (ML) methods and compare their respective likelihood values (Brueckner, 1998).

Table 1 shows the results of estimating a local tax determination equation with the level of the ordinary (business) property tax rate as the dependent variable. Apart from socio-demographic characteristics and fiscal variables, the tax determination model includes a political control dummy (that is intended to capture systematic ideological differences between right-wing and left-wing governments) and an election year dummy, that is intended to control for the fact that incumbent governments might recognize that local tax increases can be more harmful to their popularity (and to their re-election chances) close to the election date, and might therefore trim tax increases when approaching the elections and raise taxes safely after them. As already noted, in the Italian system of local government, the term of office lasts for five years (since 1999), with only a fraction of Comuni holding local elections in any given year (see table A.1, summary statistics). Since we analyze the tax setting decisions of a cross-section of Comuni for financial year 2000, we can exploit the fact that local administrators in office in year 2000 have been elected at various times in the five years 1995 to 1999, and will run again for re-election (if they do not face a term limit) in a number of years varying from 0 (if elections are to be held in 2000) to 4 (if they were elected in 1999). Consequently, we can build electoral year dummy variables, depending on how close in the future will the next elections be.

Column one of table 1 shows the OLS estimation results of a non spatial tax determination equation, while column two shows the benchmark OLS estimation results of a tax determination model where a spatially weighted average of neighboring jurisdictions' tax rates is used as an explanatory variable. The spatial weights matrix is based on a standard border-sharing criterion and is row-standardised (see note 21). Columns three and four in table 1 report the ML estimates of the two spatial models. The main results from the estimation of the key parameters show that:

1. there is no statistically significant evidence of systematic differences between right-wing and left-wing governments in property tax rate setting policies[26];
2. there is strong and significant evidence of electoral cycles: tax rates tend to be systematically lower when elections approach;
3. the spatial lag dependence model (L) yields a ML estimate of the spatial coefficient ρ of about 0.18,

[24] The Moran test on the OLS residuals was performed using a row-standardised weights matrix, based on a border-sharing criterion (see note 21 above). The test is distributed as a standard normal and takes on the value of 0.14 (z=2.7).

[25] The LM test is distributed as a $\chi^2(1)$ and takes on the value of 4.80.

[26] Solé Ollé (2001) finds evidence of more copycatting behavior in left-wing government.

that is statistically significant at the 95% level of confidence. Compared to the estimate of p in column two (OLS), it appears that OLS leads, as expected, to an upward bias in the estimation of the spatial lag parameter p. This is due to the fact that OLS is unable to discriminate between lag dependence (parameter p) and error dependence (parameter r). The increase in likelihood of model (L) with respect to a model that arbitrarily sets p=0 is statistically significant at the (not too impressing) 90% level, as shown by the likelihood ratio test (Bivand, 1984) in the bottom of table 1, third column;

4. the spatial error dependence model (E) yields a ML estimate of the spatial coefficient r of almost 0.4, that is statistically significant at the 99% level, and leads to a substantial improvement with respect to a non spatial specification that sets r=0, as shown by the likelihood ratio test (statistically significant at above 99%). The empirical evidence is consequently in favour of model (E);
5. the structural (area, population, urbanisation rate) and demographic (% young, old and unemployed) variables appear to have a significant impact on the level of the local tax rate. In particular, the size of population has a negative and highly significant impact on the tax rate, suggesting that larger municipalities might exploit economies of scale in public service provision. We also included a squared population term to allow for a non-linear relationship, but the coefficient turned out not to be significantly different from zero. On the other hand, and in accordance with most of the applied local public economics literature, both sparseness of population and urbanisation rate appear to put pressures on municipality budget constraints, while higher percentages of unemployed, elderly and young people are associated with lower tax rates;
6. the "fiscal" variables do not appear to have any systematic impact on the tax rate: all coefficients are rather imprecisely and erratically estimated. While for the income and tax base variables this could be due to large measurement error, theory itself does not univocally predict the effect of lump-sum grants on local tax rates – for instance, the existence of a 'flypaper effect' would require a very small (negative) effect of grants on the local tax rate.

From the above results, it turns out that the location of the municipalities has some relevance in explaining their taxation decisions. In particular, the positive spatial auto-correlation seems to be attributable to the fact that the unobserved components of local tax rates are correlated, as described by model (E). As we argued above, model (E) is compatible both with the presence of spatially correlated 'shocks' to tax setting behavior that have no behavioral significance, and with the sort of empirical interdependence that one should expect to find in a yardstick competition framework. According to the latter interpretation, voters in each jurisdictions should compare the tax rate in their own jurisdiction conditional on local characteristics, to neighbors' tax rates conditional on neighbors' characteristics, and penalise their own government if the 'unpredicted' component of the local tax rate is high relative to neighbors. This comparative performance evaluation after controlling for neighbors' characteristics would give rise to yardstick competition between local governments in tax rates unexplained by local characteristics. Unfortunately, such process of imitation conditional on observed characteristics is indistinguishable from the fact that common shocks to local governments are actually occurring.

To discriminate between the two alternatives, we need a further identifying condition. We therefore revert to our previous argument, exploiting the difference in the sample between 'first period' and 'last period' mayors. If the correlation in tax rates is simply due to common shocks, we should expect tax rates to be correlated among neighboring jurisdictions, whether or not the mayor faces a term limit. On the other hand, if it is yardstick competition that generates spatial auto-correlation in taxes, we should expect to find stronger correlation among those jurisdictions where mayors seek re-election. Table 2 shows the estimation results of a tax setting equation where we explicitly allow for different behavior of mayors, depending on whether they can run for re-election or they face a term limit. Again, baseline OLS estimation results of a non spatial model and of a spatial lag model are reported in the first and second columns of table 2 respectively, while the ML estimation results are reported in the third and fourth columns.

The results give in this case a clear-cut conclusion. They show that for those municipalities where the mayor steps down because of a binding term limit, there is no spatial auto-correlation in tax rates, while for those mayors that can run for re-election there is significant positive correlation in the unobservable

part of the tax setting equation. The estimated coefficient in the spatial error dependence model of column 4 in table 2 is a high and statistically significant value of almost 0.5. On the other hand, even mayors that face a term limit tend in general to keep tax rates low in the year of the elections: this may be due to party discipline. The likelihood ratio test with respect to a non spatial specification is again significant at above 99% level.

A further test of the yardstick competition hypothesis consists in exploiting the fact that while some incumbent mayors hardly passed the post in the previous election, some other mayors got an overwhelming majority of the votes. On the basis of the argument in section 2, it is reasonable to expect that the latter should feel free of implementing their preferred tax policies, irrespective of neighbors' behavior. Consequently, we build a "strong majority" dummy that equals one if the mayor got more than 50% of the votes at the first election round, and interacted it with neighbors' taxes. The results in table 3 show that there is no strategic interaction among neighboring governments for the mayors that run for re-election and are backed by very large majorities.[27] Moreover, when interacting the large majority dummy with the election year dummy, it also emerges that mayors that perceive a widespread consensus do not tend to reduce tax rates when elections get closer. In other words, unlike 'lame ducks' (that appear to be insensitive to neighbors, but accountable to the parties that support them) and mayors actively seeking re-election (that are accountable to voters and feel the competition with their neighbors), mayors who are confident of re-election appear to pursue their desired tax policies, irrespective of party discipline and neighbors' behavior.

To sum up, the results in table 2 and 3 clearly offer support to yardstick competition theory. Since positive spatial auto-correlation emerges only with regard to mayors having a sort of re-election concern, it is unlikely that it can be driven exclusively by correlated cost shocks.

5.2 Popularity equation

Turning now to the popularity equation (5), the estimates in table 4 use own and neighbors' property tax rates as explanatory variables in an equation where the dependent variable is the percentage of votes earned by incumbents at local elections.[28] All equations include the share of the vote got by the incumbent in the previous election on the right hand side, to allow for shocks to incumbents' popularity that might be persistent over time, as well as political party dummies interacted with election years, and a dummy variable that equals one if the mayor in power runs for re-election.[29] We also included as explanatory variables in the popularity equation a number of socio-economic characteristics already discussed (such as the rate of unemployment, the level of per capita income, the proportions of young and elderly people), and interacted them with political party dummies to allow for systematically different impacts of those variables on incumbents from different political parties. Since none of the estimated coefficients were statistically significant and the estimates of the parameters on the crucial variables were hardly affected, we do not report those results here.

The first column in table 4 reports the results of an OLS estimate of the popularity equation with own and neighbors' raw property tax rates on business properties on the right hand side of the equation. Notice that, considering municipalities that had elections either in 1998, or in 1999 or in 2000, our

[27] Table 3 only shows OLS estimates of a tax determination model with a spatially lagged dependent variable, that allows for different patterns of interaction for mayors that are confident of re-election and mayors that face a binding term limit. The Jacobian of the transformation between the dependent variable and the error is in this case:

$|\det(I - p_1LW - p_2F(I - L)W - p_3(I - L)(I - F)W)|$

where p_1, p_2 and p_3 are the parameters to be estimated, I is the (143×143) identity matrix, W is the (143×143) spatial weights matrix, L is the (143×143) matrix that contains information on term limits, and F is the (143×143) matrix that contains information on incumbents' confidence of re-election. When attempting to estimate this model by ML, the likelihood did not converge to a unique vector of parameters.

[28] We also estimated a popularity equation with the logarithm of the odds ratio as the dependent variable, but the results were not substantially different.

[29] The latter variable is intended to control for a potential advantage of the incumbent mayor over the opposing candidate. It must be admitted that the above variable is likely to be endogenous in the popularity equation, since the decision to run again is affected by the expected chances of success. Excluding it from the equation, though, does not affect the other coefficients considerably.

sample includes now 97 municipalities.[30] The coefficient on the lagged dependent variable (the share of the vote got in the previous election) shows a certain degree of persistence of party affiliation. Moreover, there is also evidence of mayor affiliation: when a mayor runs for a second term of office, he can expect a 6 percentage point advantage over a new candidate. On the other hand, the estimated impact of national politics (political party dummies interacted with election year dummies) on local elections is estimated to be small and not significant. As for the fiscal variables, the own property tax rate has the expected negative impact on the re-election chances of the incumbent, but the coefficient is not statistically significant. As far as neighboring effects are concerned, and contrary to the predictions of a yardstick competition model, the neighbors' tax turns out to have a negative, though not statistically significant, impact on the share of the vote of the incumbent.

In column 2, own and neighbors' taxes are instrumented with their determinants (the exogenous variables in the tax setting equation) in order to control for potential endogeneity of the fiscal variables in the popularity equation. First, we regress own and neighbors' taxes on their respective explanatory variables. Second, the predicted values of own and neighbors taxes are used as explanatory variables in the popularity equation. The coefficients on own and neighbors' taxes, though, have signs opposite to expectations, and turn out not to be statistically significant.

Finally, in column 3 we include as explanatory variables own and neighbors' unanticipated components of the tax rates, that is the residuals of the estimation of own and neighbours' tax setting equation. In this case the coefficients have the expected signs: the own unpredicted tax rate has a negative impact, while the neighbors' unpredicted tax rate has a positive impact on the popularity of the incumbent, but neither of them is statistically significant.[31]

Summing up, the electoral implications of yardstick competition theory are not corroborated by the results of the estimation of the popularity function on our data set. Note however that there are at least three main reasons which may explain the above results. First, it could be that what really matters for the popularity of the incumbent is the *change* in the tax rate, rather than its level. We consequently performed the same estimation of the popularity equation with the change in the share of the vote of the incumbent government from election t and election t+1 as dependent variable, and with own and neighbors' tax rate changes as explanatory variables. The results, though, did not show any substantial improvement. The problem however is that when tax changes are used, virtually no proper instruments are available, because most of the determinants of the tax rate in our data set (that is, the explanatory variables in the tax setting equation of section 4.1) are observed just for one year and do not change over time.[32]

Second, it must be pointed out that the low precision in the estimation of the effects of the fiscal variables on popularity could actually be the very consequence of the high degree of spatial auto-correlation between own and neighbors' tax rates. In particular, as the results in table 1 show, the residuals from the tax setting equation that are used as regressors in column 3 of table 4 are strongly and positively correlated among neighbors. Consequently, the little independent variation in the two variables – coupled with the small size of the sample – might be responsible for the high standard errors and imprecise estimates.

Finally, a third reason why the impact of property tax rates on popularity could turn out to be smaller than expected is that governments may change other variables simultaneously, such as the allocation of

[30] Whereas, as said, tax equation estimation refers to a sample of 147municipalities.

[31] A final attempt which we made is to substitute in the regressions for the popularity equation the domestic property tax rate to the ordinary one, on the basis of the argument that the domestic tax rate may be more relevant for electoral purposes as it is paid by resident households owning their own house. The results of the estimations revealed the expected strong negative effect of own domestic tax rate and a positive but insignificant effect of neighbors' tax rate on the popularity of the incumbent. It is hard however to interpret these results as evidence against or in favor of yardstick competition theory. As we argued above, in the case of the domestic property tax rate, comparing the statutory tax rates has little meaning, as municipalities apply in different ways and to a different extent the lump sum deduction allowance.

[32] In passing, this is also the reason why we preferred to specify all the models we estimated in terms of tax *levels* rather than in terms of tax *changes*, although the latter specification may have the advantage of eliminating unobservable fixed effects across jurisdictions (see Besley and Case, 1995a).

public spending among different services or public service quality. Unfortunately, no proxies of those variables are available in our data set. As regards public service quality, we might expect it to have the opposite effect on popularity with respect to tax burden, and the former to be positively correlated with the latter. Consequently, the coefficient on the included tax variables will capture as well the effect of the omitted quality variables, and will clearly turn out to be biased downwards. A deeper analysis of the simultaneous setting of multiple policy instruments, though, clearly goes beyond the scope of this work.

6. Concluding remarks

Yardstick competition is potentially one of the most interesting theories that have been proposed to explain fiscal interaction among local governments. Knowing if it is something more than just a theoretical construct is therefore important. Finding ways to test the theory, discriminating it from other competitive explanations, is not however an easy task. On the one hand, as we proved theoretically, yardstick competition theory can be consistent with different empirical predictions about tax setting behavior. On the other hand, disentangling empirically sheer spatial auto-correlation from substantive strategic interaction among neighboring jurisdictions is difficult to do, especially if one considers that some form of spatial auto-correlation among neighboring jurisdictions is one of the pre-requisite of the theory itself. To solve the problem, we looked at the key insight of the theory, the link between tax setting behavior and electoral behavior. The results are somewhat mixed. On the one hand, our tax setting model provides a rather satisfactory explanation of local tax setting determination. Moreover, by exploiting the institutional features of the Italian system of local government, we have shown that spatial auto-correlation in tax rates cannot simply be driven by correlated shocks, but it rather proves to be driven by strategic considerations, in agreement with yardstick competition theory. On the other hand, estimation results from the popularity equation do not provide clear evidence in favour of yardstick competition, and most variation in electoral outcomes remains unexplained by observable variables. Clearly, this last issue deserves more theoretical and empirical research.

Another interesting finding of our analysis is that we find clear evidence supporting the existence of electoral cycles in tax setting. Going back from empirics to theory, this suggests to look more carefully at the relationship between yardstick competition theory and the modern theory of electoral cycles, as discussed, for example, by Rogoff (1990) and Alesina *et al.* (1997). In particular, in a world which is becoming more integrated and that therefore is likely to enlarge the possibility of comparative performance evaluation across countries, it ought to be asked if yardstick competition would tend to dampen or exacerbate the electoral cycles and if there are ways to test the resulting predictions.

Finally, we must also stress the caveats of our present analysis. The most serious is that we had no way to control for the quality of public expenditure across local government, which was then assumed implicitly to be constant across jurisdictions. Finding some ways to control for quality in the empirical analysis would greatly increase our confidence in the empirical results. Again, this is a task for further research.

Table 1 Business property tax rate determination equation

	OLS non spatial model (C)	OLS spatial lag model (L)	ML spatial lag model (L)	ML spatial error model (E)
	1	2	3	4
constant	3.981*** (3.47)	3.971*** (3.00)	3.975*** (3.20)	3.974*** (3.34)
area	0.031*** (3.33)	0.028*** (3.01)	0.029*** (3.15)	0.032*** (3.42)
population	-0.523*** (3.63)	-0.500*** (3.50)	-0.508*** (3.60)	-0.580*** (3.95)
urbanisation rate	1.236*** (2.42)	1.171*** (2.32)	1.195*** (2.36)	1.433*** (2.51)
% young	-0.168*** (2.69)	-0.160*** (2.53)	-0.163*** (2.64)	-0.151*** (2.53)
% old	-0.118*** (3.01)	-0.105*** (2.66)	-0.110*** (2.85)	-0.108*** (2.72)
unemployment rate	-0.044 (1.01)	-0.049 (1.17)	-0.047 (1.13)	-0.073* (1.68)
grants per capita	0.045 (0.06)	0.221 (0.30)	0.157 (0.23)	0.576 (0.76)
income per capita	-0.037 (1.12)	-0.040 (1.22)	-0.039 (1.17)	-0.041* (1.30)
tax base per capita	-0.003 (0.94)	-0.002 (0.78)	-0.003 (0.84)	-0.001 (0.42)
left-wing party dummy	0.023 (0.23)	0.018 (0.13)	0.020 (0.22)	0.055 (0.58)
second year of term dummy	0.241** (1.83)	0.277** (2.20)	0.264** (2.03)	0.302*** (2.44)
third year of term dummy	0.067 (0.55)	0.100 (0.82)	0.088 (0.73)	0.142 (1.20)
fourth year of term dummy	0.010 (0.04)	0.054 (0.22)	0.038 (0.15)	0.196 (0.83)
last year of term dummy	-0.465** (1.90)	-0.488** (1.98)	-0.479** (2.00)	-0.551*** (2.34)
spatial parameter: p		0.281** (2.08)	0.177** (1.70)	
spatial parameter: r				0.372*** (2.95)
likelihood ratio test			2.706*	7.093***
R squared	0.251	0.274		
observations	143	143	143	143

Notes
1. dependent variable = ordinary ICI rate (business property tax rate) in year 2000;
2. absolute t statistics in parentheses;
3. p = spatial auto-correlation coefficient in dependent variable; r = spatial auto-correlation coefficient in errors;
4. spatial matrix for computing neighbors' taxes = geographical (border-sharing) and row-standardised;
5. ***=99% level of statistical significance; **=95% level of statistical significance; *=90% level of statistical significance.

Table 2 Business property tax rate determination equation: the impact of term limits

	OLS non spatial model (C)	OLS spatial lag model (L)	ML spatial lag model (L)	ML spatial error model (E)
	1	2	3	4
constant	3.987*** (3.55)	3.980*** (2.93)	3.983*** (3.44)	3.982*** (3.45)
area	0.035*** (3.73)	0.032*** (3.45)	0.033*** (3.56)	0.036*** (3.74)
population	-0.583*** (4.00)	-0.568*** (3.86)	-0.573*** (4.09)	-0.634*** (4.29)
urbanisation rate	1.344*** (2.65)	1.276*** (2.52)	1.288*** (2.58)	1.548*** (2.71)
% young	-0.188*** (2.85)	-0.191*** (2.41)	-0.196*** (2.96)	-0.178*** (2.82)
% old	-0.126*** (3.04)	-0.118*** (2.64)	-0.122*** (3.02)	-0.113*** (2.81)
unemployment rate	-0.039 (0.99)	-0.044 (0.89)	-0.044 (1.14)	-0.063* (1.50)
grants per capita	-0.102 (0.13)	0.096 (0.15)	0.042 (0.06)	0.310 (0.41)
income per capita	-0.033 (1.04)	-0.036 (1.13)	-0.036 (1.12)	-0.033 (1.10)
tax base per capita	-0.003 (0.99)	-0.003 (0.87)	-0.003 (1.05)	-0.001 (0.48)
left-wing party dummy	0.008 (0.05)	0.004 (0.04)	0.007 (0.09)	0.047 (0.48)
no term limit				
election year dummy	-0.420* (1.44)	-0.517** (1.71)	-0.495** (1.72)	-0.672*** (2.33)
spatial parameter: p		0.320** (1.92)	0.257** (1.75)	
spatial parameter: r				0.467*** (2.50)
binding term limit				
election year dummy	-0.853** (2.05)	-0.808** (1.61)	-0.820** (2.03)	-0.758** (1.91)
spatial parameter: p		0.074 (0.31)	-0.030 (0.15)	
spatial parameter: r				0.141 (0.59)
likelihood ratio test			3.071*	6.650***
observations	143	143	143	143

Notes
1. dependent variable = ordinary ICI rate (business property tax rate) in year 2000;
2. absolute t statistics in parentheses;
3. p = spatial auto-correlation coefficient in dependent variable; r = spatial auto-correlation coefficient in errors;
4. spatial matrix for computing neighbors' taxes = geographical (border-sharing) and row-standardised;
5. ***=99% level of statistical significance; **=95% level of statistical significance; *=90% level of statistical significance.

Table 3 Business property tax rate determination equation: the impact of term limits and confidence of re-election

	OLS spatial lag model (L)
constant	3.984*** (3.41)
area	0.031*** (3.44)
population	-0.558*** (4.03)
urbanisation rate	1.238*** (2.46)
% young	-0.190*** (3.03)
% old	-0.118*** (3.05)
unemployment rate	-0.048 (1.24)
grants per capita	-0.070 (0.11)
income per capita	-0.037 (1.18)
tax base per capita	-0.003 (0.93)
left-wing party dummy	-0.001 (0.09)
no term limit – confidence of re-election (share>50%)	
election year dummy	-0.167* (0.31)
neighbors' tax	0.285 (0.42)
no term limit – no confidence of re-election (share<50%)	
election year dummy	-0.734** (1.85)
neighbors' tax	0.328** (1.93)
term limit	
election year dummy	-0.802** (2.01)
neighbors' tax	0.082 (0.34)
observations	143

Table 4 Popularity equation: impact of own and neighbors' property tax rates

	1	2	3
constant	25.591 (1.02)	31.244 (0.85)	17.404*** (2.72)
share of the vote [t-1]	0.456*** (4.02)	0.450*** (4.00)	0.454*** (4.02)
dummy (same mayor runs = incumbency advantage)	6.390*** (2.35)	6.041** (2.19)	6.195*** (2.24)
dummy (right-wing coalition)	3.848 (0.57)	3.493 (0.53)	4.204 (0.64)
dummy (left-wing coalition)	8.326* (1.45)	7.523* (1.32)	8.519* (1.50)
dummy (right-wing coalition & election year 1999)	2.350 (0.26)	4.237 (0.46)	3.002 (0.33)
dummy (right-wing coalition & election year 2000)	3.249 (0.26)	6.828 (0.53)	3.895 (0.311)
dummy (left-wing coalition & election year 1999)	-3.830 (0.90)	-2.214 (0.49)	-3.655 (0.86)
dummy (left-wing coalition & election year 2000)	-5.830 (0.79)	-2.736 (0.36)	-5.653 (0.78)
own business property tax	-0.744 (0.36)		
neighbors' business property tax	-0.627 (0.18)		
predicted own business property tax		4.444 (0.89)	
predicted neighbors' business property tax		-6.839 (1.04)	
unpredicted own business property tax			-1.252 (0.60)
unpredicted neighbors' business property tax			1.088 (0.32)
R squared	0.234	0.243	0.236
observations	97	97	97

Notes

1. dependent variable = share of the vote got by the incumbent government at elections held at a given election year (election year equals either 1998, or 1999, or 2000);
2. absolute t statistics in parentheses;
3. spatial matrix for computing neighbors' taxes = geographical (border-sharing) and row-standardised;
4. ***=99% level of statistical significance; **=95% level of statistical significance; *=90% level of statistical significance.

Acknowledgements

Previous versions of this paper have been discussed in seminars held at the University of Rome III, the University of Milan Bicocca, the University of Brescia, the SIEP Conference in Pavia and at an International Conference organized by MEP at the Catholic University of Milan. We wish to thank all participants for the useful comments and in particular Jan Brueckner, Rosella Levaggi, Ben Lockwood, Piergiovanna Natale, Fabio Padovano, Jorn Rattso and Patrizio Tirelli. Usual disclaimers apply. We are also grateful to Aldo Vaccaro from the Home Ministry for providing data on political variables. Floriana Cerniglia gratefully acknowledges financial support from CNR 1999, contr. N. 99.01644 CT10.

Appendix 1

Table A.1 Summary statistics

VARIABLE	OBS.	MEAN	STANDARD DEVIATION	MIN	MAX
ordinary ICI tax rate (‰)	143	5.905	0.625	4	7
domestic ICI tax rate (‰)	143	5.150	0.529	4	6
area	143	1150.112	1579.958	198	18158
population	143	25,269	110,327	4,005	1,321,555
urbanisation rate	143	18.408	15.325	1.808	85.180
percentage young	143	21.225	1.981	15.1	29
percentage old	143	16.585	3.232	5.9	25.1
percentage unemployed	143	7.840	1.677	4.751	13.603
business tax base	143	69.767	17.540	1.042	96.599
domestic tax base	143	71.750	7.665	32.670	84.062
grants per head (million liras)	143	0.323	0.095	0.057	0.856
income per head (million liras)	143	13.694	2.143	10.086	23.967
political dummy (right)	143	0.091	0.276	0	1
political dummy (centre)	143	0.361	0.481	0	1
political dummy (left)	143	0.548	0.499	0	1
election in 1996	143	0.042	0.201	0	1
election in 1997	143	0.203	0.403	0	1
election in 1998	143	0.175	0.381	0	1
election in 1999	143	0.538	0.500	0	1
election in 2000	143	0.042	0.201	0	1
share of the vote	97	44.61	11.58	21.3	72.9
same mayor runs	97	0.762	0.427	0	1

References

Alesina, A., Roubini, N., and Cohen, G. (1997) *Political Cycles and Macroeconomics*, MIT Press, Cambridge.

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

Anselin, L., Bera, K., Florax, R., Yoon, M. (1996) "Simple diagnostic tests for spatial dependence," *Regional Science and Urban Economics*, **26**, 77-104.

Anselin, L., Florax, R., eds., (1995) *New Directions in Spatial Econometrics*, Springler-Verlog, Heidelberg.

Besley, T., Case, A. (1995a) "Incumbent behavior: vote seeking, tax setting and yardstick competition," *American Economic Review,* **85**, 25-45.

Besley, T., Case, A. (1995b) "Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits," *Quarterly Journal of Economics* ,**110**, 769-798.

Besley, T., Smart, M. (2001) "Globalization and Electoral Accountability," unpublished paper.

Bivand, R. (1984) "Regression modeling with spatial dependence: an application of some class selection and estimation methods," *Geographical Analysis,* **16**, 25-37.

Bivand, R., Szymanski, S. (1997) "Spatial dependence through local yardstick competition: theory and testing," *Economics Letters,* **55**, 257-265.

Bivand, R., Szymanski, S. (2000) "Modelling the spatial impact of the introduction of Compulsive Competitive Tendering," *Regional Science and Urban Economics,* **30**, 203-219.

Bordignon, M., Minelli, E. (2001a) "Rules transparency and political accountability," *Journal of Public Economics,* **80**, 73-98.

Bordignon, M., Minelli, E. (2001b) "Tax competition and yardstick competition," unpublished paper.

Breton, A. (1998) *Competitive Governments*, Cambridge University Press, Cambridge.

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

Brueckner, J. (1998) "Testing for strategic interaction among local governments: the case of growth controls," *Journal of Urban Economics*, **44**, 438-467.

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

Buettner, T. (2001) "Local business taxation and competition for capital: the choice of tax rate," *Regional Science and Urban Economics*, **31**, 215-245.

Case, A. (1993) "Interstate tax competition after TRA86," *Journal of Policy Analysis and Management*, **12**, 136-148.

Case, A., Hines, J., Rosen, H. (1993) "Budget spillovers and fiscal policy interdependence," *Journal of Public Economics,* **52**, 285-307.

Coate, S. and Morris, S. (1995) "On the forms of transfers to special interests," *Journal of Political Economy*, **105**, 1210-1235. .

Figlio, D., Kolpin, V., Reid, W. (1999) "Do states play welfare games?" *Journal of Urban Economics,* **46**, 437-454.

Heyndels, B., Vuchelen, J. (1998) "Tax mimicking among Belgian municipalities," *National Tax Journal,* **60**, 89-101.

Kelejian, H., Robinson, D. (1993) "A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model," *Papers in Regional Science,* **72**, 297-312.

Ladd, H. (1992) "Mimicking of local tax burdens among neighboring counties," *Public Finance Quarterly,* **20**, 450-467.

Meyer, M. and Vickers, J. (1997) "Performance comparisons and dynamic incentives," *Journal of Political Economy*, **105**, 547-581.

Murdoch, J., Rahmatian, M., Thayer, M. (1993) "A spatially autoregressive median voter model of recreation expenditures," *Public Finance Quarterly*, **21**, 334-350.

Oates, W. E. (1972) *Fiscal Federalism*, Harcourt Brace Jovanovich, New York.

Paldam, M. (1997) "Political business cycles," in D. Mueller, (ed.), *Perspectives on Public Choice: a Handbook,* Cambridge University Press, Cambridge, 342-370.

Persson, T. and Tabellini, G. (2000) *Political Economics, Explaining Economic Policy*, MIT Press.

Revelli, F. (2001a) "Local taxes, national politics and spatial interactions in English district election results," *European Journal of Political Economy* (forthcoming).

Revelli, F. (2001b) "Spatial patterns in local taxation: tax mimicking or error mimicking?" *Applied Economics*, **33**, 1101-1107.

Rogoff, K. (1990) "Equilibrium political budget cycles," *American Economic Review*, **80**, 21-36.

Saavedra, L. (2001) "A model of welfare competition with evidence from AFDC," *Journal of Urban Economics*, **47**, 248-279.

Salmon, P. (1987) "Decentralization as an incentive scheme," *Oxford Review of Economic Policy*, **3**, 24-43.

Shroder, M. (1995) "Games the states don't play: welfare benefits and the theory of fiscal federalism," *The Review of Economics and Statistics*, **77**, 183-191.

A. Solè Ollè, (2001) "Tax mimicking and electoral control: an empirical analysis of local tax setting in Spain", Universitat del Barcelona, *mimeo.*

Tiebout, C. (1956) "A pure theory of local expenditures," *Journal of Political Economy,* **64**, 416-424.

Wildasin, D. (1986) *Urban Public Finance*, Harwood Academic Publishers, New York.

Wilson, J. (1999) "Theories of tax competition," *National Tax Journal,* **53**, 269-304.

Wrede, M. (2001) "Yardstick competition to tame the Leviathan," *European Journal of Political Economy*, **17**, 705-721