Yardstick Competition and Partial Coordination: Exploring the Empirical Distribution of Local Business Tax Rates

Thiess Buettner (FAU and CESifo)†

Axel von Schwerin (FAU)

May 2015‡

Abstract:

This paper explores a striking form of tax policy interdependence that can be observed in the German federation. Though municipalities enjoy discretion in setting the local business tax rate, large fractions of municipalities – in some states even the majority – set identical tax rates. Our analysis shows that this tax-rate "bunching" is not the result of federal or state-level institutions. Possible explanations rest on partial coordination and yardstick competition. The role of the former is exemplified by the finding that small jurisdictions and jurisdictions sharing the same county are more likely to engage in "bunching". Yardstick competition seems also relevant since jurisdictions for which strategic tax-setting should be associated with larger gains and lower cost are in fact more likely to set identical tax rates.

Keywords: Local business tax; Tax competition; Yardstick competition; Partial tax coordination; Location; Asymmetric information

JEL classification: H71; H77

†Address: Friedrich-Alexander-University
Lange Gasse 20
D-90403 Nuremberg
Germany

Phone: +49 911 5302 200
Fax: +49 911 5302 396
E-mail: thiess.buettner@fau.de

‡For helpful comments on an earlier draft we are indebted to special issue guest editor Veronika Grimm, two anonymous referees, Bev Dahlby, Simon Jaeger, Stefan Traub, and seminar participants on various occasions.

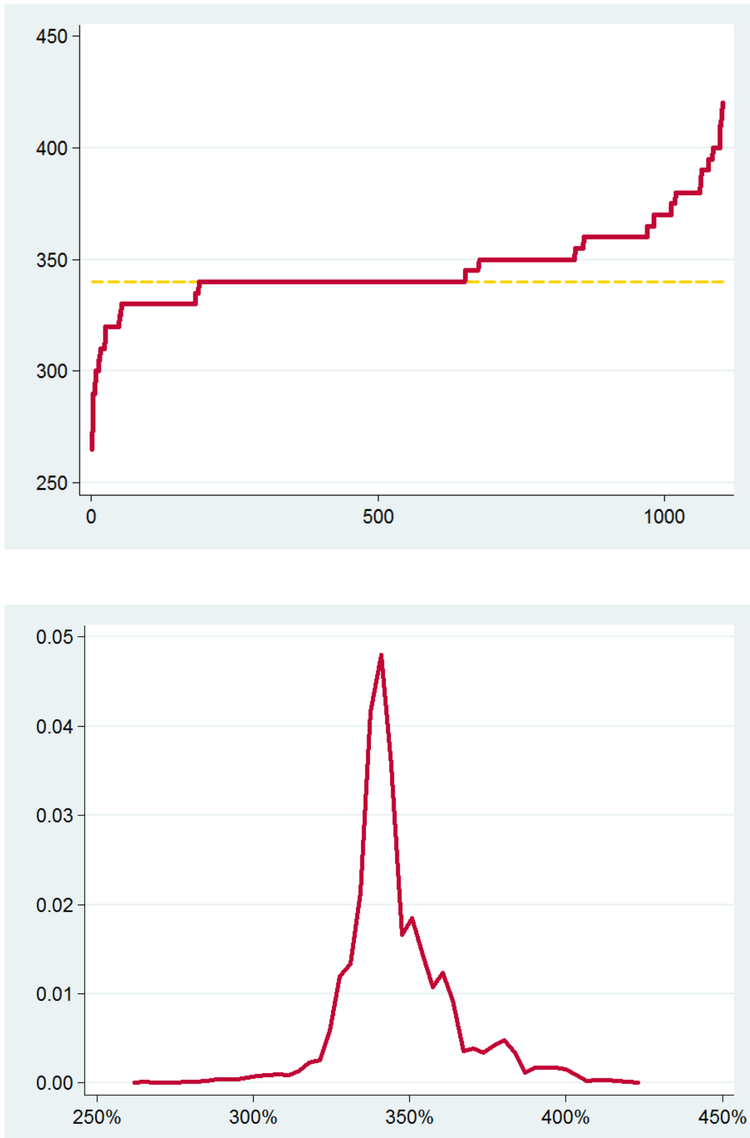
1 Introduction

The literature on fiscal decentralization has emphasized that decentralization substantially alters the conditions for public policy. A large body of literature has focused on competition for mobile factors and noted that if the set of available fiscal instruments is limited, local fiscal policies including taxation decisions become interdependent (*e.g.*, Wilson, 1986, and Zodrow and Mieszkowski, 1986). If jurisdictions pursue non-cooperative policies, coordination fails and the resulting equilibrium is inefficient. As central government intervention is required to restore efficiency (*e.g.*, Wildasin, 1989), decentralization appears to be self-defeating. Another strand of the literature has emphasized that decentralized policies open up the opportunity to compare local policies and their outcomes between jurisdictions. The role of comparative performance information (Meyer and Vickers, 1996) in a setting with decentralized policies is featured in particular in the literature on yardstick competition (*e.g.*, Besley and Case, 1995, Wrede, 2001, Bordignon, Cerniglia, Revelli, 2004). The basic argument is that, since economic shocks are correlated, comparisons between jurisdictions are meaningful, and voters engage in comparative policy evaluation, which limits the ability of the governments to extract rents.

The empirical literature has provided ample evidence that the local fiscal choices such as local tax policies are in fact interdependent. Many papers have considered capital taxation in the presence of mobile investors, and found evidence for spatial interdependence in local property or business taxes (*e.g.*, Buettner, 2001, Brueckner and Saavedra, 2001, Bordignon, Cerniglia, Revelli, 2003) as well as in taxation of FDI (*e.g.*, Devereux, Lockwood, Redoano, 2008). The empirical literature also shows that governments eager to get re-elected, tend to respond stronger to neighbors' tax policies (*e.g.*, Besley and Case, 1995, Revelli, 2002, Bordignon, Cerniglia, Revelli, 2003). However, it has proved difficult to verify whether interdependence results from comparative performance evaluation or from mobility (Revelli, 2005).

An interesting form of tax policy interdependence can be studied in the German federation, where municipalities enjoy the right to set the tax rate of the local business tax (Gewerbesteuer) – a tax on the earnings of corporations as well as partnerships. The local business tax is the single most important own revenue source of municipalities in the German federation. With statutory rates between about 7% and 18%, the tax rate of the local business tax is a key parameter of the effective tax rate. Since investors are confronted with important differences in the tax burden and also risk to be taxed at high rates, the business tax multiplier of each municipality is widely documented and reported by employers associations, municipalities, statistical offices and the media. Empirical research shows that the elasticity of the tax base is large (Buettner, 2003) and that even the location decisions of multinational firms respond to the local tax rate (Becker, Egger, Merlo, 2012). Although local jurisdictions in Germany enjoy autonomy in setting the local business tax rate, empirical evidence shows that large numbers of municipalities deliberately choose identical tax rates. Figure 1 illustrates this phenomenon using data for the about 1100 municipalities in the state of Baden-Wuerttemberg. The figure depicts the distribution of the local multiplier of the basic business tax rate in percent. For instance, if the local municipality sets the multiplier to a level of 400%, the resulting statutory business tax rate amounts to 14% (= 400% $\times$ 0.035). The upper panel plots the business tax multiplier of the municipalities ordered by size. There are 464 municipalities, almost half of all municipalities in this state, that have chosen a business tax multiplier of exactly 340 percent. The kernel density plot in the lower panel highlights the striking density mass at this tax rate.

Figure 1: Distribution of the Local Business Tax Multiplier



The figures depict the cross-sectional distribution of local business tax rates among the total population of 1101 municipalities in Baden-Wuerttemberg in the year 2011. The figures report the tax multiplier which determines the statutory business tax rate. For instance, with a multiplier of 400 percent the resulting statutory business tax rate amounts to 0.14 or 14% (= 400% $\times$ 0.035). The top panel displays the business tax multiplier of all municipalities in percent in ascending order (red, solid line) and the mode of the distribution (yellow, dashed line). The lower panel provides a corresponding kernel density plot of the distribution.

In Baden-Wuerttemberg as in the other German states, municipalities are small villages or big cities, some of which are located in rural – others in metropolitan areas, and there is usually a large diversity in economic conditions and cultural traditions. Since the jurisdictions differ so much, it seems reasonable to expect differences in preferences and economic conditions to be reflected in a non-trivial distribution of tax rates. Due to tax competition, it would not be surprising if the tax rates are found only within a certain range. However, the example of Baden-Wuerttemberg shows that there is excessive "bunching" at specific tax rates (*cf.* Saez, 2010, for the notion of "bunching"). The important point here is that this is not the result of any regulation.

In this paper we explore this striking form of tax policy interdependence. We review approaches taken in the literature to explain local tax policy decisions and discuss which of these approaches offer possible explanations for tax-rate "bunching". We take a broader perspective and consider a standard tax competition setting, but also yardstick competition and related forms of signaling as well as institutions and partial coordination and highlight differences in the empirical implications of these approaches. Backed with this discussion we provide econometric tests as to whether the deliberate decision of a large number of jurisdictions to set identical tax rates is a consistent pattern among German municipalities. Finally, we explore the characteristics of jurisdictions that take part in tax-rate bunching and test whether some of the theoretical predictions can be confirmed.

The results support tax-rate bunching as a common phenomenon among German municipalities across states and over time. As standard forms of tax competition and institutions can be ruled out as explanations, yardstick competition and partial coordination turn out to be the most convincing explanations.

The remainder of this paper is organized as follows. The next section considers various theoretical explanations for bunching. Section three discusses different ways to test empirical implications of the different explanations. In section 4 we get back to the empirical distribution of business tax rates set by local jurisdictions in Germany and provide descriptive statistics by state. Section 5 provides empirical results on the issue whether bunching at specific points of the tax rate distribution is a consistent pattern among German municipalities, whether it is a permanent or transitory phenomenon and whether certain institutional explanations can be ruled out. The characteristics of jurisdictions that engage in tax-rate bunching are explored in section 6. Section 7 provides a brief summary and concludes.

2 Theoretical Explanations of Tax-Rate Bunching

Bunching of tax rates in the sense of a large fraction of jurisdictions setting identical tax rates can be rationalized using various theoretical approaches to tax policy, albeit in some cases only under highly restrictive assumptions. In this section, we discuss different approaches that can be found in the theoretical literature.

Tax Competition The textbook model in public finance explains the choice of the tax rate in terms of optimal provision of public services. In this approach, pioneered by Samuelson (1954) and Musgrave (1959), the provision of public services exerts benefits but public funding is costly. As the marginal cost of public funds is increasing in the tax rate, the optimal tax rate ensures that the marginal benefit from public service provision is equated with the marginal cost of funds. This framework also applies in a decentralized setting. In the literature on tax competition the standard model (*e.g.*, Zodrow and Mieskowski, 1986, Wilson, 1986) discusses a tax on a mobile factor, where the marginal cost of funds faced by the individual jurisdiction reflects the adverse effect of a higher tax burden on the location of capital. With higher taxes, for example, other jurisdictions benefit from the inflow of factors and experience a fiscal externality (Wildasin, 1994). As a consequence, tax policy decisions become interdependent. In a symmetric setting with a large number of identical jurisdictions, the Nash-equilibrium would result in identical tax rates. However, this is an extreme result that does not hold if jurisdictions differ. Any asymmetries between jurisdictions such as differences in jurisdiction size or in preferences will give rise to a non-trivial tax rate distribution. From this point of view, bunching of tax rates in the sense of setting identical tax rates should be rather unlikely among heterogeneous jurisdictions. As the Nash equilibrium might not be the most appropriate characterization of the tax competition equilibrium between asymmetric jurisdictions, the literature has considered Stackelberg equilibria, where large jurisdictions move first and set tax rates, while others follow (*e.g.*, Fredriksson and Millimet, 2002, Altshuler and Goodspeed, 2014). While this might cause jurisdictions to set their tax rate close to the tax rate of the Stackelberg leader, still a non-trivial tax distribution should be expected.

Coordinating Institutions Various explanations for tax-rate bunching in the sense of setting identical tax rates are associated with the gains that could be reaped by tax coordination. Since the tax competition equilibrium is inefficient, jurisdictions have an incentive to cooperate and coordinate their tax policy. However, the literature has stressed the difficulties of jurisdictions to commit to a cooperative solution (for an excellent survey see Keen and Konrad, 2013). Hence, much attention has been paid to coordination enforced by some institution. Harmonization (*e.g.*, Sinn, 1990, Razin and Sadka, 1991), for instance, involves identical tax rates and thus is consistent with some trivial form of bunching. In the presence of asymmetries, however, harmonization is less likely to occur since it will not equally benefit all jurisdictions and may even harm some of them (Kanbur and Keen, 1993). Another institutional form of coordination is to define a minimum tax rate. Even if the minimum tax rate is not binding in the sense that in the initial equilibrium all jurisdictions set higher tax rates, theoretical research suggests that defining a floor in a tax competition game might give rise to bunching at the minimum tax-rate (Konrad, 2009).

Partial Coordination Coordination might also take place without enforcing institutions. While most of the theoretical literature on tax competition deals with one-shot games, interjurisdictional competition might be better characterized as a (infinite) repeated game. In such a setting coordination is obtained if jurisdictions are sufficiently patient and if asymmetries are limited (*e.g.*, Cardarelli, Taugourdeau, Vidal, 2002, Catenaro and Vidal 2006). This suggests that partial tax coordination may emerge between subgroups of more symmetric jurisdictions. The theoretical literature has further discussed conditions under which tax competition equilibria with partial coordination are more likely to emerge and has emphasized that successful coordination requires tax rates to be strategic complements (Burbidge, DePater, Myers and Sengupta, 1997, Konrad and Schjelderup, 1999). In an asymmetric setting, small jurisdictions might find partial tax coordination more beneficial than large jurisdictions (Vrijburg and deMooij, 2010). However, as jurisdictions joining as well as those staying apart benefit from coordination, it is not always clear which jurisdictions might be joining. As is known from cooperative game theory, coordination might be supported by implicit threats exercised upon a violation of the agreement (Gordon, 1992). As local jurisdictions within a region interact in various dimensions, this would suggest that coordination is more likely to arise between adjacent jurisdictions or jurisdictions within certain geographical distances or in specific regions.

Heuristics and Herding Given the complexity of identifying beneficial strategies in a setting with policy interdependence it might be rather difficult for a group of jurisdictions to establish tax coordination. Tax-rate bunching might reflect the presence of heuristics that are featured in the literature on focal points and prominent numbers (*e.g.*, Keser and Vogt, 2000, Pope, Selten, Kube, 2009). As a simple form of heuristics, jurisdictions might be choosing a business tax multiplier that is a "round number" such as 340% compared to rates of 341% or 338%. As we discuss below, another form of heuristics might be to choose business tax multipliers equal to reference points provided by upper-level governments. However, heuristics might not necessarily be associated with coordination, since they could simply reflect the lack of government information about the consequences of tax decisions.[1] Lacking information, governments themselves might also tend to utilize information about other jurisdictions' tax policy decisions when choosing their tax rate. If tax policy choices are taken sequentially, some form of herding (Banerjee, 1992) might occur, where jurisdictions just follow choices of others. While it seems intuitive to argue that tax policy choices of governments work as signals to other governments, each jurisdiction can also observe previous own tax choices as well as choices of other jurisdictions and their consequences and, hence, learning might be important (Becker and Davis, 2015). Moreover, in the presence of asymmetries, it is not clear why herding should result in identical tax rates. Herding might simply be confined to follow trends in taxation such as common intellectual trends (Devereux, Lockwood, Redoano, 2008).

Yardstick Competition In the presence of information asymmetries with regard to the objectives of local governments, an alternative explanation for bunching is some form of yardstick

[1] Taking account of psychological factors, also some practice of "odd taxation" might be observed. In this fashion, Olsen (2013) studies local income tax rates of Danish municipalities, where municipalities set income tax rates some digits below round numbers, such as 19.8% or 19.9% instead of 20%.

competition. The literature has focused in particular on the lack of information of voters who compare local policies with policies of other jurisdictions. This opens up the strategic option for local governments to mimic policies of other governments in order to provide a signal to voters. As shown by Besley and Case (1995) and Bordignon, Cerniglia, Revelli (2004), this could give rise to a pooling equilibrium where local governments set identical tax rates. Since the tax policy of jurisdictions that share similar economic conditions is most informative, tax mimicking should be most relevant among jurisdictions that share basic characteristics. Those might include population size such that jurisdictions of the same size are more likely to display identical tax rates. Also the geographic location might be an important common characteristic. This supports the view that bunching might be most relevant between neighboring jurisdictions. Of course, geographic proximity might foster tax mimicking also due to information cost, if information is more easily gathered in nearby places. Consequently, the empirical literature on yardstick competition has tested wether tax mimicking is more prevalent within local media markets (Revelli, 2008).

The standard theory of yardstick competition emphasizes that comparative performance evaluation is important for voters who have the choice to vote for or against a government. In a context of local jurisdictions mobility offers also a choice where to locate. In the presence of adjustment cost that makes it difficult to relocate once decisions are made, governments cannot commit to low tax rates in the future and investors face a risk of expropriation (*e.g.*, Thomas and Worrall, 1994). Hence, investors will find it crucial to carefully study not only the basic location characteristics but also to engage in a comparative policy evaluation of local jurisdictions. With imperfect information, investors will consider current policy choices to form beliefs on future policy. As we show in the appendix using a stylized formal model, this could also give rise to some form of yardstick competition between local jurisdictions. As the model shows, concerned about the effects of their current policy on the expectations about future tax policy, jurisdictions evaluate strategies to deviate from their preferred tax policy, in order to provide a signal to investors that the risk of a high tax rate in the future is low. We derive conditions under which this setting gives rise to a pooling equilibrium where local jurisdictions mimic each others' policies. More specifically, even if investors anticipate that the local tax policy could be reversed in the future, under certain conditions the dominant strategy for some jurisdictions is to set tax rates equal to the median of the observed tax distribution, even though the preferred tax-rate is higher.[2] Intuitively, we might expect, in particular, jurisdictions that attach a high value to raising investment to rely on signaling. In the formal analysis these jurisdictions are modeled as jurisdictions that aim at increasing wages of the immobile factor. Possible drivers behind this objective could be a relatively small share of old residents who live on pensions rather than labor income or a large share of young residents. Also high levels of unemployment might explain why jurisdictions put more emphasis on attracting investors.

Though the theoretical analysis abstracts from public infrastructure provision and debt finance, a possible extension is a setting where public infrastructure is provided. In order to keep providing infrastructure services to investors, a low-tax strategy to attract investors might require to run deficits. If the debt level is already high, jurisdictions might find that it is costly and difficult to set low tax rates. High debt might also undermine the credibility of a low-tax strategy. Hence, we might expect that jurisdictions with high levels of debt are less likely to follow a bunching strategy.

[2] For analytical convenience the theoretical model employs a simple two-period setting. A more realistic model would have $n + 1$ periods, where the first period with mobile capital is subdivided into n periods and new investors may enter a jurisdiction in all of these sub-periods. In such a setting jurisdictions may still deviate from their tax preference and pursue a low-tax strategy for multiple sub-periods until a sufficient amount of capital is accumulated.

3 Empirical Implications

The theoretical analysis has discussed various reasons why a group of heterogeneous jurisdictions might set identical tax rates including coordinating institutions, partial coordination, and yardstick competition. To see whether tax-rate bunching is a consistent pattern in the tax policy decisions of local jurisdictions in Germany and whether and how well it fits with the various explanations, we explore the tax rate distribution in a number of different ways. A first issue is whether bunching in the sense of multiple jurisdictions setting identical tax rates is a common phenomenon among German municipalities and to what extent it reflects institutions. A second issue is how tax-rate bunching evolves over time. A third issue explored below is whether there are some common characteristics of those jurisdictions that have chosen to set identical tax rates.

3.1 Significance of Tax-Rate Bunching

As we have discussed above, the distribution of tax-rates in a tax competition equilibrium would reflect differences in preferences and in the cost of funds. However, if preferences and cost of funds differ, tax-rate bunching seems unlikely - except if there are coordinating institutions. In the context of the business tax there are in fact some institutions that might give rise to bunching of tax rates. First of all, a minimum tax rate that constrains jurisdictions in their tax policy choices could be important. This is relevant in the German context studied below, as a minimum tax rate for the business tax was introduced in 2004. If many jurisdictions set a tax-rate equal to the minimum rate, bunching is not necessarily the result of coordination or tax mimicking but might simply reflect constrained optimization. Another potential explanation in the German case comes from the federal income tax. The business tax paid to the local municipality can be deducted from the federal income tax, at least for partnerships. The federal income tax, however, only accepts a deduction of taxes up to a certain threshold level of the business tax multiplier. More specifically, since 2008 the income tax grants a deduction of the business tax burden up to a multiplier of 380%. This might induce jurisdictions to raise their tax rate exactly to this level (see Buettner, Scheffler, von Schwerin, 2014). Another important institutional explanation in the German context is offered by state-specific grant programs, which distribute funds among municipalities according to some measure of the tax capacity of each jurisdiction. In order to compute the tax capacity some states define specific reference rates for the business tax multiplier – other states rely on the average tax rate.[3] While the average tax rate adjusts from year to year, the specific reference rate chosen by a number of states is changed only occasionally. In some states such as Baden-Wuerttemberg it has remained unchanged over decades. This specific reference rate might well serve as a focal point in the tax coordination among local jurisdictions. However, its impact on local tax choices could also be driven by the political cost of raising taxes. If a jurisdiction sets a tax rate equal to the reference rate, the local government could try to sell its choice by some sort of neutrality argument: the local businesses are taxed at a rate that the state's grant system considers to be the standard tax effort. In this case, bunching would not indicate interdependence in taxation but simply reflect a setting where for a couple of jurisdictions the political assessment of taxing options is characterized by a joint optimum.

To control for similarities in preferences and cost of funds, round numbers or specific institutions

[3]The latter procedure is similar to the representative tax system approach which was introduced to compute the tax capacity of Canadian provinces and has also been used in the US revenue sharing system which was abolished in 1987 (see Courchene, 1984, Wildasin, 1987).

that apply in the German context we employ an empirical approach by Chetty, Friedman, Olsen and Pistaferri (2011).[4] More specifically, we model the count of municipalities that choose a specific tax rate and explore whether the count of municipalities choosing a tax rate equal to the mode of the tax rate distribution is significantly higher than what any continuous distribution of preferences or cost of funds would predict. To this end we build bins defined by specific values of the local business tax multiplier, and explore a regression

$$\begin{aligned} C_k &= \alpha_0 + \sum_{j=1}^{p} \alpha_{1,j} m_k^j + \alpha_2 I\left(m_k \in \{200\%, 210\%, ...\}\right) \\ &+ \alpha_3 I\left(m_k = minimum\right) + \alpha_4 I\left(m_k = federal\right) + \alpha_5 I\left(m_k = capacity\right) \\ &+ \beta I\left(m_k = mode\right) + u_k. \end{aligned} \tag{1}$$

With $k \in \{200\%, 201\%, 202\%, ...\}$ indicating specific choices of business tax multipliers, C_k denotes the count of all municipalities that choose multiplier k. To capture the effects of continuous distributions in preferences and cost of funds, we include a higher-order polynomial of the specific tax multiplier m_k of each count as a set of controls. A couple of further controls check wether some institutional characteristic or the existence of some simple heuristics might explain the tax-rate bunching. $I\left(m_k \in \{200\%, 210\%, ...\}\right)$ captures tax multipliers that are multiples of 10 to test for round (prominent) number effects, $I\left(m_k = minimum\right)$ captures the minimum multiplier (200%), $I\left(m_k = federal\right)$ captures the multiplier used to compute the tax deduction for business taxes in the federal income tax (380%), $I\left(m_k = capacity\right)$ captures a fixed reference level for the multiplier used by the state's grant program to define tax capacity, if available. Finally, we add a dummy variable capturing specifically the mode of the distribution of the business tax multipli-

[4] Chetty *et al.*, 2011, consider a quite different form of bunching that refers to tax filing by individuals and which arises at kink points of the income tax code.

ers $I(m_k = mode)$. If $\beta > 0$, we can confirm that bunching at this rate cannot be explained by similarities in continuously distributed preferences or cost of taxation among jurisdictions, or by institutions and simple heuristics.

The analysis below provides evidence for the state of Baden-Wuerttemberg, which has already been featured in the introduction. We also provide results for other states. Since the number of municipalities varies strongly by state, slope parameters will tend to vary by state. Hence, pooling of municipalities among states is not an option. Instead, the analysis below provides separate regressions for individual states. In each of these regressions we employ a state-specific mode of the tax rate-distribution as well as fixed reference values of the multiplier used to define the tax capacity in the state's grant program.

3.2 Persistence of Tax-Rate Bunching

In the standard model of tax competition, in the absence of adjustment costs, an unexpected change in preferences or in the economic conditions faced by a jurisdiction would be accompanied with a change in tax policy that restores equilibrium. If jurisdictions have access to debt finance, this needs to be qualified: due to tax smoothing (Barro, 1979), only permanent changes would trigger a noticeable adjustment. However, if permanent shocks to preferences or economic conditions occur randomly for all jurisdictions, we should expect to see changes all over the tax rate distribution. With cost of relocation and cost of migration the adjustment will not be instantaneous but be extended over a period of time, as capital stocks, labor and tax policies are slowly adjusting (see Wildasin, 2003). Given that the speed of adjustment is a function of the adjustment cost, all else equal, we should expect that the adjustment of the tax rate distribution takes place gradually in all jurisdictions.

Under coordination of tax policies, a permanent perturbation in the economic conditions faced by a single jurisdiction and permanent changes in a single jurisdiction's preferences might have no impact on the coordinated tax rate. This would suggest that bunching due to tax coordination should display a higher persistence. Also with yardstick competition, changes in preferences and economic conditions will not necessarily trigger changes in tax policy. Only with regard to common shocks immediate responses should be expected.[5]

To test for the persistence of bunching we explore whether municipalities that have set their tax rate equal to the mode show a lower likelihood to change their tax rate. We use a framework that takes account of the possibility that the changes in preferences or cost of funds vary somehow along the tax rate distribution. This also helps to control for some regression to the mean effects. More specifically, we test for differences in the likelihood of changing the business tax rate using a specification which involves controls for the current level of the tax rate. Formally,

$$
\begin{aligned}
I\left(\Delta m_i \neq 0\right) &= \gamma_0 + \sum_{j=1}^{p} \gamma_{1,j} m_i^j + \gamma_2 I\left(m_i \in \{200\%, 210\%, ...\}\right) \\
&+ \gamma_3 I\left(m_i = minimum\right) + \gamma_4 I\left(m_i = federal\right) + \gamma_5 I\left(m_i = capacity\right) \\
&+ \delta I\left(m_i = mode\right) + u_i.
\end{aligned} \tag{2}
$$

The dependent variable $I\left(\Delta m_i \neq 0\right)$ is a binary indicator of whether the local tax multiplier of jurisdiction i is changed in the following time-period. To capture the observed distribution of

[5]If yardstick competition is driven by imperfect information on the part of the voters, also re-elections should trigger tax rate changes, at least if the voting rules imply that this is the last term of the government (*e.g.*, Besley and Case, 1995).

tax preferences, we include a higher-order polynomial of the specific tax multiplier chosen in the current period m_i as controls. Besides controls for institutional effects or tax-rate heuristics, we add a dummy variable $I(m_i = mode)$ capturing specifically municipalities that have the tax rate set equal to the mode. If $\delta < 0$, we can confirm that municipalities with tax rate equal to the mode display a lower propensity to change their tax rate.

The analysis considers the evidence for the state of Baden-Wuerttemberg and we also provide results for other states. As above, we employ state-specific values for the modes of the tax rate-distribution and for the reference rate of the multiplier, if a fixed rate is used to define the tax capacity in the state's grant program.

3.3 Determinants of Tax-Rate Bunching

A third issue is to explore the characteristics of individual jurisdictions that set identical tax-rates and to test whether theoretical predictions can be confirmed. In order to test the various potential determinants of bunching we focus on the tax rate distribution in one German state and utilize a linear probability model to see whether the likelihood to follow a bunching strategy is significantly correlated with a jurisdiction's characteristics. Formally,

$$I(m_i = mode) = \theta_0 + \theta x_i' + u_i. \tag{3}$$

The dependent variable $I(m_i = mode)$ is a binary indicator of whether the local tax multiplier of jurisdiction i is set equal to the mode of the empirical distribution. x_i' is a vector of characteristics of jurisdiction i that captures potential determinants of tax-rate bunching.

The above theoretical discussion has suggested that if bunching in the sense of a large number of jurisdictions setting identical tax rates is not associated with simple heuristics or specific institutions, it could reflect partial coordination as well as yardstick competition. The theoretical literature suggests that coordination is difficult in the presence of asymmetries in jurisdictions' size. Under certain conditions, coordination in tax competition is more likely among smaller jurisdictions as they can reap larger gains from the coordination. Since asymmetries in preferences and cost of funds tend to be generally associated with more unequal gains from coordination, one might also hypothesize that coordination could be more easily established among jurisdictions that have equal preferences or economic conditions. This suggests that not only population size but also that the composition of the population in terms of the age structure and, at least for some regions in Germany, the religious affiliation (Catholic Church, Lutheran Church, *etc.*) could help to predict bunching. The likelihood to engage in coordination might also be correlated with certain political preferences: jurisdictions where a certain political party has won the majority might be more often engaged in tax coordination. This could reflect similarities in ideology which might help to establish coordination due to a similar business orientation. Also hierarchical patterns might be important, in the sense that coordination is more likely if the same party holds the majority in the state government. Another determinant of the likelihood to engage in coordination might be the number of seats in the municipal council. As the number of mandates jumps discontinuously with increasing population size, there is some variation in the number of mandates per capita.[6] If local politics suffer from a common pool problem (*e.g.,* Weingast *et al.*, 1981) a jurisdiction with a larger number of legislators might find it more difficult to engage in tax coordination, since the benefits from tax coordination are shared among a larger number of agents.

[6]See Egger and Koethenbuerger (2010), for an analysis of discontinuities in the election law for Bavarian municipalities.

A general problem with tax coordination in tax competition games is that individual jurisdictions might benefit from defecting. If coordination is the driving force, we might expect bunching to be more prevalent in settings where individual jurisdictions interact not only in terms of taxes but also in terms of other dimensions of policy, as this opens up the possibility for implicit threats exercised upon a violation of the agreement. From this perspective, we should expect that tax coordination is most easily established at the local level between neighboring jurisdictions or between jurisdictions in the same region. If interaction takes place within certain institutions such as some upper-level jurisdiction (region, province) that defines certain policies for a group of jurisdictions, bunching as some form of coordination might also be more likely among jurisdictions that belong to the same province or region. In the German context, the local jurisdictions (municipalities) form counties. Counties provide some public services and infrastructure to all municipalities. They also have a budget to which municipalities pay contributions. It seems likely, therefore, that the likelihood to establish tax coordination might be higher among jurisdictions that belong to the same county.

Neighborhood and spatial effects are, of course, also an implication of yardstick competition. In the literature, yardstick competition is usually associated with a setting where voters face uncertainty about government objectives and governments strive for re-election. Voters then engage in comparative performance evaluation using the nearby jurisdictions as reference points. However, as we have discussed above, yardstick competition also arises in a context of mobile investors that face uncertainty about future tax policy. In the stylized model sketched above, jurisdictions that are concerned about the future wages of the immobile factor, have an incentive to deviate strategically from their tax preference and to mimic tax policies of others including, of course, neighboring jurisdictions. This kind of argument might be quite important in the context of the German business tax as the tax base is highly mobile. From the above discussion, a signaling strategy is more likely to be perceived as beneficial by a jurisdiction that aims at increasing investment. Empirically, this suggests that bunching is more likely in jurisdictions that face higher unemployment and host many young people and have less old residents. As we have noted above, a low level of public debt might also raise the likelihood of bunching as it makes a low-tax strategy both more credible and less difficult/costly for the local government.

4 Data and Descriptive Statistics

To explore tax-rate setting among German municipalities we collected data on local business tax multipliers in all 13 German non-city states in 2011.[7] Table 1 provides some information about the mode, *i.e.* the most frequently chosen tax multiplier in each state. As is documented in the table, in most states a large fraction of municipalities choose identical tax rates. The smallest figure is observed for Saarland, which however comprises only 52 municipalities. The largest fraction of bunching municipalities is observed for Mecklenburg, where 403 out of 805 municipalities set a multiplier of 300%.

In some states (Lower-Saxony, Saxony, and Saarland) the most frequently chosen business tax multiplier is equal to the reference rate used by the federal income tax (380%). In these states, it might be difficult to rule out an institutional explanation for bunching. Moreover, in some states that use a fixed multiplier to define tax capacity (see column 6), the most frequently chosen multiplier is equal to this rate. However, in Baden-Wuerttemberg, the case highlighted in the

[7]The city states of Berlin, Hamburg and Bremen do not have separate municipalities.

Table 1: Modes and Reference Rates of the Local Business Multiplier

State	State Pop. (in 1000) (1)	No. of Munic. (2)	Avg. Pop. Size (3)	Mode of Tax Multi-plier (4)	No. of Munic. with tax equal mode (5)	Reference rates for Tax-Capacity (6)	Federal Inc. Tax (7)
Baden-Wuerttemberg	10,769	1,101	9,781	340	464	290	380
Bavaria	12,560	2,056	6,109	320	386	300	380
Brandenburg	2,498	419	5,963	300	144	.	380
Hessen	6,075	426	14,260	310	82	310	380
Mecklenburg	1,638	805	2,034	300	403	.	380
Lower Saxony	7,926	1,010	7,847	380	143	.	380
North-Rhine Westfalia	17,893	396	45,185	403	58	403 [1)]/411	380
Rhineland-Palatinate	3,999	2,306	1,734	352	569	352	380
Saarland	1,015	52	19,514	380	7	.	380
Saxony	4,139	470	8,806	380	110	.	380
Saxony Anhalt	2,323	220	10,558	300	40	.	380
Schleswig-Holstein	2,835	1,116	2,541	310	292	310	380
Thuringia	2,227	913	2,439	357	249	300/357 [2)]	380

State characteristics in 2011. (1) State population in thousands. (2) Number of municipalities. (3) Average population size of municipalities. (4) Mode of tax multiplier denotes the most frequently chosen business tax multiplier in the state. The tax multiplier determines the statutory business tax rate. For instance, with a multiplier of 400 percent the resulting statutory business tax rate amounts to 0.14 or 14% (= 400% $\times$ 0.035). (5) Number of municipalities with multiplier equal to (4). (6) Reference rate used to define tax capacity for the state specific equalization scheme; missing values reflect states that rely on averages. (7) Reference rate used to define the federal income tax deduction. [1)]: reference rate applicable until 2010. [2)]: new reference rate announced in 2010.

introduction, the system of equalization grants also employs a fixed multiplier to define tax capacity, which is, however, much lower (290%) than the mode. In Bavaria, the mode (320%) is also different from the fixed multiplier (300%).

In order to avoid dealing with cases, where bunching may just reflect some institutionally fixed reference points, the subsequent analysis focuses only on states where the mode of the tax rate distribution is not equal to some fixed reference rate used to define tax capacity or the tax deduction for the federal income tax. Thus we focus on the tax multiplier chosen by the municipalities in the five states of Baden-Wuerttemberg, Bavaria, Brandenburg, Mecklenburg and Saxony-Anhalt.

5 Significance and Persistence of Tax-Rate Bunching

To explore whether the number of jurisdictions that set the same tax rate is significantly higher than what any continuous distribution of preferences or cost of funds would suggest, we employ the regression approach outlined above. Table 2 shows results from various specifications that estimate the regression equation (1) using data for the state of Baden-Wuerttemberg. The dependent variable is the count of municipalities in each bin defined by a specific value of the local tax multiplier. Since the tax law requires the municipalities to define the business tax multiplier in terms of percentage points, we consider 316 bins starting from the minimum tax multiplier of 200 and up to a maximum multiplier of 515.[8]

The constant in the first column shows that on average three municipalities share a common business tax multiplier. However, as shown in column (2), about 464 municipalities (sum of the constant and the slope parameter) out of the total number of 1101 have set exactly the same tax multiplier. We can also see this number reflected in column (5) of table 1. But regardless of whether or not we use more or less complex polynomials of the tax rate to capture an arbitrary tax rate distribution, this tax rate is chosen by many more municipalities than any continuous distribution would suggest. Columns (3) to (7) take account of various institutional explanations for bunching, by identifying bins where the tax multiplier is equal to some reference figure such as the minimum tax multiplier, a fixed multiplier used to compute tax capacity, and the reference figure for the multiplier defined

[8]While the minimum rate in Baden-Wuerttemberg in 2011 is 265 and the maximum is 420, the range of bins used here reflects a wider distribution. The minimum rate is defined by federal legislation. The maximum tax rate is not legally defined, but except for a few tiny municipalities in Rhineland-Palatinate that charge extremely high multipliers, the highest multiplier is observed in the state of Northrhine-Westfalia with a level of 515. Hence, the range of 316 bins from 200 to 515 enables us to cover comprehensively all the different state-specific distributions.

Table 2: Tax-Rate Bunching in the State of Baden-Wuerttemberg

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
I (mode)		462.0 ***	446.0 ***	445.9 ***	443.6 ***	443.4 ***	441.3 ***
		(0.784)	(8.053)	(8.091)	(8.708)	(8.795)	(9.379)
I (minimum)			-18.04 **	-18.51 **	-13.94 **	-12.85 **	-18.94 **
			(8.053)	(8.282)	(6.669)	(6.284)	(8.239)
I (capacity)			-14.04 *	-14.25 *	-15.38 *	-15.87 *	-15.08 *
			(8.053)	(8.156)	(8.365)	(8.594)	(8.160)
I (federal)			24.96 ***	25.02 ***	22.72 ***	22.94 ***	20.99 **
			(8.053)	(8.042)	(8.643)	(8.554)	(9.092)
I (div.by 10)			17.74 **	17.75 **	17.85 **	17.86 **	18.06 **
			(8.054)	(8.069)	(7.937)	(7.942)	(7.867)
m_k				-0.290	19.79 **	45.28	-640.5 **
				(0.375)	(7.677)	(28.82)	(266.6)
m_k^2					-2.805 **	-10.21	294.5 **
					(1.088)	(7.705)	(121.8)
m_k^3						0.690	-57.65 **
						(0.667)	(23.80)
m_k^4							4.074 **
							(1.682)
Constant	3.484 **	2.022 **	0.299 ***	1.337	-32.27 **	-60.24 *	500.1 **
	(1.657)	(0.784)	(0.114)	(1.350)	(12.67)	(34.40)	(210.0)
R^2	0.000	0.778	0.814	0.814	0.819	0.819	0.823
R^2 corr.	0.000	0.778	0.811	0.810	0.815	0.814	0.818
Number of obs.	316	316	316	316	316	316	316

Dependent variable: count of municipalities that set a specific local tax multiplier. The count refers to the empirical distribution of tax multipliers among all 1101 municipalities in the year 2011 covering a range from 200% to 515%. This results in 316 observations. I (mode) characterizes the mode of the distribution of tax multipliers (340%). I (minimum) denotes the minimum tax multipliers (200%). I (capacity) identifies the reference rate used to define fiscal capacity (290%). I (federal) captures the reference rate used by the federal income tax (380%). I (div. by 10) refers to all multipliers that are multiples of 10 (200%, 210%, 220%,...). m_k captures the business tax multiplier identifying the bin. OLS results. Robust standard errors are given in parentheses. * denotes significance at 10%, ** significance at 5%, *** significance at 1% level.

by the federal income tax. We also include an indicator for bins where the business tax multiplier is a multiple of 10. According to the specification reported in column (7), about 440 municipalities more than expected under a general non-parametric model have chosen a business tax multiplier equal to the mode (340%). There is also some evidence for bunching at the reference rate used in the federal income tax as well as a higher propensity to set the multiplier equal to some multiple of 10. The appendix shows separate results for other states (see table A-1). While reference rates defined by federal and state legislation exert some significant effects, the results are quite similar to those obtained for Baden-Wuerttemberg and strongly support tax-rate bunching in all states.

Results for tests of tax-rate persistence among municipalities in the state of Baden-Wuerttemberg are presented in table 3. The dependent variable is an indicator of whether the tax rate in 2011 differs from 2009.[9] The constant in column (1) shows that about 20 percent of all municipalities set a different tax rate in 2011 relative to 2009. Columns (2) to (4) test whether municipalities which have set their tax rate equal to the mode of the distribution in 2009 show a significantly lower likelihood to adjust their tax rate. Regardless of whether controls for the tax rate are included, the mode shows a negative significant effect indicating that the likelihood to change the tax rate is substantially lower for municipalities that have set their tax rate equal to the mode. Note that an indicator for the minimum multiplier is omitted since no municipality in this state has set its multiplier at the minimum. Regression results for the other states are provided in the appendix (see table A-2). The results are somewhat mixed. In Mecklenburg, the state with largest share of municipalities that set their business tax multiplier equal to the mode, a significantly lower likelihood to change the tax rate is confirmed for these municipalities. This is also confirmed for

[9]We took a period of two years rather than one to raise the likelihood to observe changes. A longer period is not used, since in 2008 major changes in the tax system were enacted that limit comparability.

Table 3: Likelihood of a Tax Rate Change in Baden-Wuerttemberg

	(1)	(2)	(3)	(4)
I (mode)		-0.116 ***	-0.117 ***	-0.146 ***
		(0.023)	(0.024)	(0.028)
I (div.by 10)			-0.001	-0.031
			(0.051)	(0.050)
I (federal)			-0.036	0.128
			(0.096)	(0.105)
I (capacity)			0.004	0.066
			(0.218)	(0.235)
m_i				0.556
				(1.572)
m_i^2				-0.002
				(0.007)
m_i^3				0.000
				(0.000)
m_i^4				-0.000
				(0.000)
Constant	0.196 ***	0.245 ***	0.247 ***	-59.81
	(0.012)	(0.017)	(0.047)	(132.3)
R^2	0.000	0.021	0.021	0.039
R^2 corr.	0.000	0.020	0.017	0.032
Number of obs.	1101	1101	1101	1101

Dependent variable: Binary variable indicating whether the tax rate set by jurisdiction i differs in 2011 from 2009. Controls refer to 2009. I (mode) is a binary variable indicating jurisdictions with tax multiplier equal to the mode of the distribution of tax multipliers (340%). I (capacity) is a binary variable indicating jurisdictions with tax multiplier equal to the reference rate used to define fiscal capacity (290%). I (federal) is a binary variable indicating jurisdictions with tax multiplier equal to the reference rate used by the federal income tax (380%). I (div. by 10) is a binary variable indicating jurisdictions with tax multiplier equal to a multiple of 10 (200%, 210%, 220%,...). m_i is the business tax multiplier set by jurisdiction i. OLS regression results. Observations are 1101 municipalities in Baden-Wuerttemberg. Robust standard errors in parentheses. * denotes significance at 10%, ** significance at 5%, *** significance at 1% level.

Bavaria. For the other two states no significant effect is found. Yet as the number of jurisdictions is substantially smaller in these states also the power of the test is lower in these cases.

6 Determinants of Tax-Rate Bunching

To explore the characteristics of jurisdictions that set their tax rate equal to the mode we employ various indicators capturing the jurisdictions' characteristics. We focus on the state of Baden-Wuerttemberg, where we have been able to augment the dataset on tax rates with various municipality characteristics. Descriptive statistics are provided in table 4. The list of explanatory variables includes indicators of the population size and the population structure, as well as public debt per capita, a binary variable indicating zero public debt and unemployment per capita. Four variables capture the political dimension. This includes the number of mandates per capita as defined by the election laws of the state, and the variables indicating whether one of the dominant parties have the majority (Christian Democrats and *Freie Waehler*). As some smaller municipalities use simple majority voting without party lists, we include a dummy for these municipalities. We also use indicators of the prevalence of bunching or the absence of bunching in other municipalities weighted with inverse geographic distances with regard to each individual jurisdiction. An alternative measure of the importance of bunching within the region is the (unweighted) share of all municipalities within the county that have set their tax rate equal to the bunching point.

Basic cross-sectional regression results for the linear probability model are presented in table 5. The first two columns show effects of population size, a simple log in column (1) and a fourth-order polynomial in column (2). The log, showing a slightly better fit, indicates that bunching is more

Table 4: Descriptive Statistics for Baden-Wuerttemberg

Variable	Mean	Std.Dev.	Min.	Max.
I (mode)	0.422	0.494	0	1
W I (mode) [a)]	0.287	0.084	0.028	0.498
W I (off mode) [a)]	0.395	0.130	0.025	0.670
County average of I(mode) [a)]	0.432	0.229	0	0.880
Urban county	0.008	0.090	0	1
Population	9,787	26,521	96.00	609,256
Young population, per capita	0.211	0.020	0.145	0.299
Old population, per capita	0.189	0.028	0.107	0.363
Catholic, per capita	0.433	0.218	0.026	0.897
Protestant, per capita	0.330	0.174	0.032	0.807
Public debt (in Euro)	0.788	0.692	0	4.580
I(debt= 0)	0.083	0.276	0	1
Unemployed, per capita	0.027	0.009	0.006	0.070
I(majority vote)	0.061	0.239	0	1
Mandates, per capita	0.004	0.053	0.001	0.078
I(CDU)	0.069	0.254	0	1
I(*Freie Waehler*)	0.351	0.477	0	1

I (mode) is a binary variable indicating whether a jurisdiction has set tax rate equal to the mode. W I (mode) is the spatial average of the binary indicator of whether tax rate is equal to the mode. W I (off mode) is a spatial average of the binary indicator of whether tax rate is not equal to the mode. I (debt= 0) is a binary variable indicating zero public debt. I (majority vote) is a binary variable indicating majority voting. I (CDU), I(*Freie Waehler*) are binary variables indicating that the respective party holds the majority in the municipal council. See appendix for further definitions and sources. Due to missing values statistics refer to 1100 municipalities in Baden-Wuerttemberg in 2011.
[a)]: refers to 2010

prevalent with small jurisdictions. This result is confirmed also when further controls are included. Columns (3) and (4) include indicators of the age structure of the resident population, indicators of religious affiliation, and indicators of public debt. The latter include the total amount of public debt per capita as well as a dummy variable capturing municipalities with zero public debt. The results suggest that bunching is more likely if the resident population is relatively young, if there is stronger Christian affiliation of the resident population and if the level of public debt is low. Columns (5) and (6) add a couple of further variables which are insignificant. These include an indicator of unemployment, variables capturing election outcomes as well as number of mandates.

The results provided in table 5 are consistent with explanations of tax-rate bunching that consider it a form of partial tax-coordination. Smaller jurisdictions might be more likely to cooperate since gains from tax coordination are larger. They might tend to be less heterogenous such that gains from coordination are more equal and, hence, coordination is more easily established. Tax-rate bunching also turns out to be more likely among jurisdictions with a stronger Christian affiliation – interestingly, this holds regardless of confession. This result is in accordance with partial tax coordination, as the churches might foster links between municipalities. The religious affiliation might also capture the joint history which could make it easier to cooperate as municipalities share characteristics.

The results are also in accordance with yardstick competition/signaling as an explanation for setting identical tax rates. If smaller municipalities are less heterogenous, comparisons are more meaningful, and, hence, yardstick competition becomes more relevant. If we are willing to assume that a municipality with a younger population is more eager to attract investment, the age structure effects are also consistent with a tax-policy that engages in signaling. Also the negative significant

Table 5: Determinants of Tax-Rate Bunching: Basic Results

	(1)	(2)	(3)	(4)	(5)	(6)
Log population	-0.132 ***		-0.060 ***		-0.068 ***	
	(0.011)		(0.017)		(0.023)	
Population		-0.023 ***		-0.011 ***		-0.010 ***
		(0.002)		(0.003)		(0.003)
Population^2		0.000 ***		0.000 ***		0.000 **
		(0.000)		(0.000)		(0.000)
Population^3		-0.000 ***		-0.000 ***		-0.000 **
		(0.000)		(0.000)		(0.000)
Population^4		0.000 ***		0.000 ***		0.000 **
		(0.000)		(0.000)		(0.000)
Young pop			2.142 **	2.191 **	1.941 *	2.010 **
			(0.968)	(1.003)	(0.996)	(1.016)
Old pop			-0.500	-0.477	-0.532	-0.467
			(0.668)	(0.700)	(0.686)	(0.703)
Catholic			0.873 ***	0.875 ***	0.839 ***	0.827 ***
			(0.229)	(0.224)	(0.243)	(0.243)
Protestant			0.559 **	0.541 **	0.512 *	0.481 *
			(0.273)	(0.274)	(0.285)	(0.289)
Debt			-0.073 ***	-0.068 ***	-0.071 ***	-0.069 ***
			(0.022)	(0.022)	(0.022)	(0.022)
I (debt=0)			-0.060	-0.031	-0.050	-0.040
			(0.057)	(0.056)	(0.059)	(0.060)
Unemployed					1.133	1.432
					(1.977)	(2.008)
I (majority vote)					0.034	0.055
					(0.079)	(0.080)
Mandates					-3.758	0.722
					(4.768)	(4.356)
I (CDU)					0.102	0.097
					(0.063)	(0.063)
I (*Freie Waehler*)					0.056	0.063 *
					(0.038)	(0.038)
Constant	1.541 ***	0.571 ***	0.073	-0.380	0.174	-0.385
	(0.097)	(0.021)	(0.413)	(0.350)	(0.462)	(0.364)
R^2	0.085	0.082	0.144	0.146	0.148	0.150
R^2 corr.	0.084	0.079	0.138	0.138	0.139	0.138
Number of obs.	1100	1100	1100	1100	1100	1100

Dependent variable: Binary variable indicating whether the local tax multiplier of jurisdiction i is equal to the mode of the tax rate distribution. OLS regression results. Observations are 1100 municipalities in Baden-Wuerttemberg in 2011. Robust standard errors are given in parentheses. * denotes significant at 10%, ** significant at 5%, *** significant at 1%.

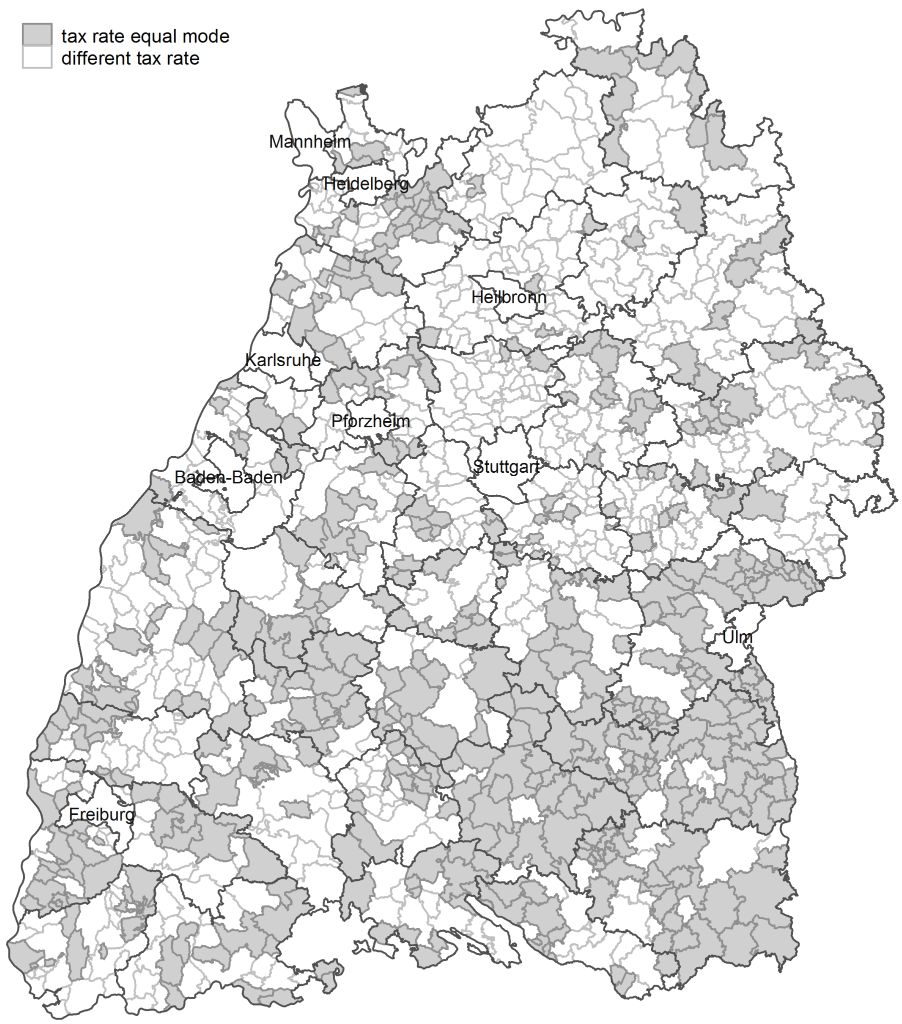
effect of public debt is in accordance with this explanation. With more public debt, municipalities may find it more costly to attract investors through imitating a low-tax policy – this policy might also be less credible. Even the significant effects of the religious affiliation might be associated with a stronger preference for using the tax rate as device to attracting investors. Since if the Christian affiliation of the constituency is stronger, also the parish's budget is larger and the cost of engaging in signaling might be smaller if some public services are provided by the local parish.

As we have discussed above, both explanations for bunching, partial coordination as well as yardstick competition, point to some spatial association between jurisdictions that set equal tax rates. Figure 2 shows that municipalities with tax rates set equal to the mode tend to cluster forming small regions with equal tax rates, in particular in the south–western part of the state. This might reflect some spatial interaction in local tax policies which arises if yardstick competition is more intense among geographically close jurisdictions or if coordination is more likely to hold between nearby jurisdictions.

To test for a spatial pattern in the linear-probability model we follow Hautsch and Klotz (2003) who study spatial effects on discrete outcomes by conditioning on spatial averages of the two outcomes of the limited dependent variable in the preceding period.[10] They also include a lagged dependent variable to control for autocorrelation. The advantage of this approach relative to spatial probit modeling is that it conditions on the realization of certain policies rather than trying to establish a spatial relationship in the latent variable (see also Dubin, 1995). To apply this method in our setting, we compute for each municipality weighted averages of binary indicators of whether or not neighboring jurisdictions have set their tax rate equal to the mode. The first has value one for all

[10] Note that the spatial averages of the outcomes are not complements.

Figure 2: Spatial Distribution of Tax-Rate "Bunching" in Baden-Wuerttemberg



The figure depicts 1100 municipalities in the state of Baden-Wuerttemberg. Municipality borders are shown in a light grey color, county borders are shown in a dark grey color. All municipalities with a tax multiplier set equal to the mode of the distribution are shaded grey.

municipalities where the tax multiplier in the previous period has been identical to the mode (340%), the second has value one if the tax multiplier has been different from the mode. Following standard practice in spatial econometrics, the averages are computed using inverse geographic distances to the respective municipalities as weights. However, we do not impose row standardization, as this would imply that spatial effects have the same strength regardless of the number of neighboring jurisdictions.

Suppressing the coefficient estimates for the control variables, table 6 displays the corresponding results for the six basic specifications shown in table 5. As the results show, to choose a business tax multiplier equal to the mode is significantly more (less) likely if neighboring municipalities have (not) chosen the same tax multiplier in the previous period. Note that this effect cannot be attributed to a combination of omitted variables and autocorrelation in the time dimension as the specifications also condition on whether the tax rate of each jurisdiction had been equal to the mode in the previous period. While the significant effects point to spatial dependence in bunching, the significance rests on the assumption that the spatial structure including the distance weights is correctly specified. It should be noted, however, that neighboring municipalities tend to share the same county. Thus the spatial effects might simply stem from an interaction within counties. This interaction may help to establish tax coordination as each county has an own budget to which the individual municipalities contribute and which may be used to share gains from coordination. One might also argue that common county policies make comparisons with jurisdictions in the same county more informative. This would suggest that interactions in the same county are also consistent with yardstick competition. As a test of whether spatial effects are equivalent to county effects, table 7 reports results from specifications where the spatial averages of the dependent

Table 6: Determinants of Tax-Rate Bunching: Tests for Spatial Effects

	(1)	(2)	(3)	(4)	(5)	(6)
I (mode)$_{t-1}$	0.821 ***	0.810 ***	0.809 ***	0.807 ***	0.808 ***	0.807 ***
	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
W I (mode)$_{t-1}$	0.498 ***	0.462 ***	0.361 ***	0.384 ***	0.379 ***	0.387 ***
	(0.105)	(0.104)	(0.110)	(0.112)	(0.116)	(0.118)
W I (off mode)$_{t-1}$	-0.304 ***	-0.276 ***	-0.213 **	-0.229 ***	-0.232 ***	-0.236 ***
	(0.072)	(0.070)	(0.084)	(0.088)	(0.089)	(0.090)
R^2	0.741	0.743	0.746	0.746	0.747	0.747
R^2 corr.	0.740	0.742	0.743	0.743	0.743	0.743
Number of obs.	1100	1100	1100	1100	1100	1100

Dependent variable: Binary variable indicating whether the local tax multiplier of jurisdiction i is equal to the mode of the tax rate distribution. I (mode)$_{t-1}=1$ is the lagged dependent variable. W I (mode)$_{t-1}$ is the spatial average of the binary indicator of whether tax rate is equal to the mode, W I (off mode)$_{t-1}$ is a spatial average of the binary indicator of whether tax rate is not equal to the mode. Both variables refer to the previous year. OLS regression results. Specifications (1) to (6) include the same controls as specifications (1) to (6) in table 5. Robust standard errors are given in parentheses. * denotes significance at 10%, ** significance at 5%, *** significance at 1% level. Observations are 1100 municipalities in Baden-Wuerttemberg in 2011.

Table 7: Determinants of Tax-Rate Bunching: Tests for County Effects

	(1)	(2)	(3)	(4)	(5)	(6)
I (mode)$_{t-1}$	0.818 ***	0.804 ***	0.804 ***	0.802 ***	0.803 ***	0.802 ***
	(0.021)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Cnty. av. of I (mode)$_{t-1}$	0.172 ***	0.168 ***	0.129 ***	0.136 ***	0.130 ***	0.131 ***
	(0.040)	(0.039)	(0.042)	(0.043)	(0.043)	(0.044)
Urban county dummy	0.006	0.040	0.082 ***	0.055	0.100 ***	0.065
	(0.014)	(0.049)	(0.028)	(0.050)	(0.036)	(0.049)
R^2	0.740	0.743	0.746	0.746	0.747	0.747
R^2 corr.	0.739	0.742	0.743	0.743	0.743	0.742
Number of obs.	1100	1100	1100	1100	1100	1100

Dependent variable: Binary variable indicating whether the local tax multiplier of jurisdiction i is equal to the mode of the tax rate distribution. I (mode)$_{t-1}=1$ is the lagged dependent variable. County average of I (mode) is an unweighted average of the binary indicator of whether tax rate has been equal to the mode taken in all municipalities of the respective county in the previous period. OLS regression results. Specifications (1) to (6) include the same controls as specifications (1) to (6) in table 5. Robust standard errors are given in parentheses. * denotes significance at 10%, ** significance at 5%, *** significance at 1% level. Observations are 1100 municipalities in Baden-Wuerttemberg in 2011.

variable are replaced by the share of municipalities in the same county that have set a tax rate equal to the mode.[11] Since there are some urban counties, where this effect is not present, the specifications include a dummy for such urban counties. Qualitatively, the findings are similar and also the fit is quite similar to the spatial specifications from table 6. This indicates that county effects and spatial effects are empirically indistinguishable.

In order to control for county effects in a general fashion we remove those effects by means of county-level fixed effects. Table 8 shows corresponding results. Qualitatively, the results are similar to those obtained above. Columns (1) and (2) show that small municipalities are more likely to set the tax multiplier equal to the mode. As documented in columns (3) and (4) the age structure suggests that jurisdictions with a larger share of older residents relative to the county average tend to be less likely to set the tax rate equal to the mode of the tax rate distribution. Again, a higher fraction of the resident population with Christian affiliation (Catholic Church and Protestant Churches) is significantly related with more bunching. Also the negative effect of public debt is confirmed. Columns (5) and (6) also support an effect of the number of unemployed, indicating that jurisdictions are more likely to set their tax rate equal to the mode if unemployment is high. If those municipalities have a higher preference for attracting investment, this is in line with the predictions from yardstick competition. In column (5) the number of mandates proves significant. However, this effect vanishes in column (6) where a polynomial of population size is employed.

[11] Since the share of municipalities that set a tax rate different from the mode is the complement, the latter indicator is not included.

Table 8: Determinants of Tax-Rate Bunching: Results with County Fixed-Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Log population	-0.101 ***		-0.056 ***		-0.108 ***	
	(0.013)		(0.018)		(0.024)	
Population		-0.018 **		-0.006		-0.013
		(0.007)		(0.008)		(0.009)
Population^2		0.000		-0.000		-0.000
		(0.000)		(0.000)		(0.000)
Population^3		0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)
Population^4		-0.000		-0.000		-0.000
		(0.000)		(0.000)		(0.000)
Young population			0.049	-0.067	-0.040	-0.018
			(0.964)	(0.967)	(0.965)	(0.977)
Old population			-1.249 *	-1.242 *	-1.409 **	-1.370 **
			(0.679)	(0.679)	(0.680)	(0.685)
Public debt			-0.048 **	-0.037	-0.048 **	-0.044 *
			(0.022)	(0.023)	(0.023)	(0.023)
I (debt=0)			-0.102 *	-0.074	-0.077	-0.059
			(0.055)	(0.054)	(0.057)	(0.057)
Catholic			0.967 ***	0.846 ***	1.148 ***	1.113 ***
			(0.311)	(0.308)	(0.328)	(0.331)
Protestant			0.988 ***	0.828 **	1.178 ***	1.122 ***
			(0.347)	(0.348)	(0.364)	(0.368)
Unemployed					3.947 *	4.507 **
					(2.025)	(2.044)
I (majority vote)					-0.050	-0.017
					(0.072)	(0.073)
Mandates					-10.27 ***	-3.254
					(3.180)	(3.035)
I (CDU)					0.062	0.053
					(0.056)	(0.056)
I (*Freie Waehler*)					-0.017	-0.005
					(0.038)	(0.039)
Constant	1.340 ***	2941	0.494	3875	0.971 **	3015
	(0.172)	(2333)	(0.417)	(2464)	(0.467)	(2507)
R^2	0.264	0.277	0.283	0.290	0.291	0.294
R^2 corr.	0.234	0.244	0.248	0.254	0.254	0.255
Number of obs.	1100	1100	1100	1100	1100	1100

Dependent variable: Binary variable indicating whether the local tax multiplier of jurisdiction i is equal to the mode of the tax rate distribution. OLS regression results. Coefficients for county-specific fixed effects are suppressed. Robust standard errors are given in parentheses. * denotes significant at 10%, ** significant at 5%, *** significant at 1%. Observations are 1100 municipalities in Baden-Wuerttemberg in 2011.

7 Summary and Conclusions

Decentralization of the public sector is controversially discussed in the literature as it gives rise to interdependency in local policies. A large body of research emphasizes the resulting inefficiency of local public policies if coordination fails. Another strand of the literature emphasizes the efficiency potential associated with comparative policy evaluation. Ample empirical evidence has been provided in the literature documenting that local fiscal policies, local taxes in particular, are interdependent. However, so far little empirical evidence is available on the extent to which coordination failure prevails and how relevant comparative policy evaluation is.

To shed light on these issues, this paper has explored a striking form of tax policy interdependence that can be observed in the German federation. Though municipalities enjoy discretion in setting the local business tax rate, a large number, in some states even the majority, deliberately chooses identical tax rates. To explain this pattern we have reviewed approaches taken in the literature to explain local tax policy decisions and have discussed which of these approaches offer possible explanations for tax-rate "bunching". While it is difficult to explain this pattern using standard tax competition models, an obvious possible explanation would be some form of institutionalized coordination. Though there are several relevant institutions in the German context, such as a minimum tax rate, the federal income tax and state specific grant programs, our empirical analysis has shown that the deliberate move to set identical tax rates is not the result of such institutions.

One promising explanation rests on the implications of comparative information for local tax policies. While the literature on yardstick competition has employed this concept in settings where voters engage in policy comparisons we have argued that comparative information may not only be important for voting decisions but also for investment decisions. If investors face adjustment costs which make it difficult to relocate once decisions are made, they will also compare policies across jurisdictions to form expectations about the local jurisdictions' future policies. With asymmetric information they will be receptive to signals, and, as a consequence, local jurisdictions may mimic each others' policies. More specifically, some local jurisdictions may hide their preference for high tax rates and rather set a tax rate equal to some common level such as the median to be similar or even indistinguishable from other jurisdictions. Another promising explanation is some form of partial tax coordination that takes place among a subgroup of jurisdictions. The empirical pattern of tax rate adjustment, pointing to a lower likelihood to observe changes if the tax rate is set equal to the mode, is consistent with either of these explanations. Under coordination, if individual jurisdictions face shocks that provide an incentive to alter tax policies, they are less likely to respond, sticking to the coordinated tax rate. Also under yardstick competition persistence should be higher, as jurisdictions that engage in tax mimicking might not adjust their tax rates to local shocks.

To test whether jurisdictions that should be more likely to engage in tax-rate bunching caused by partial coordination or yardstick competition are actually doing so, we have further explored municipal tax setting in a German state. We find that small jurisdictions are more likely to set their tax rate equal to the mode as well as jurisdictions with a higher share of young and a lower share of old population. Tax-rate bunching also turns out to be more likely among jurisdictions with a stronger Christian affiliation – interestingly, this holds regardless of the specific confession. We find evidence for strong spatial correlation which is however not distinguishable from county effects. Controlling for county effects, size, demographics and religious affiliation effects prove robust, and also jurisdictions with low level of public debt and with high unemployment are found to be more likely to follow a strategy of tax-rate bunching. Government ideology, measured by the party affiliation of the majority in the municipal county, does not show significant effects.

These results are consistent with the view that bunching is a form of partial tax-coordination. In particular, smaller jurisdictions might be more likely to cooperate since gains from tax coordination are larger. They might also tend to be less heterogenous such that gains from coordination are more equal and, hence, coordination is more easily established. Also the spatial interaction effects are consistent with coordination between municipalities within the same region or county. However, the results are also in accordance with yardstick competition and signaling as driving forces behind bunching. If smaller municipalities are less heterogenous, comparisons across municipalities are more meaningful. If we are willing to assume that a municipality with younger (older) population is more (less) eager to attract investment, the age structure effects are also consistent with signaling. Also the positive effect of unemployment and the negative significant effect of public debt is in accordance with this explanation. Higher unemployment is likely to be associated with a stronger preference for attracting investors, and with more public debt, municipalities may find it more costly to attract investors through imitating a low-tax policy – this policy might also be less credible.

Regardless of whether partial coordination or yardstick competition or a combination of both are at work, our results indicate that the strong pessimism with regard to decentralized taxation which is much emphasized in the literature on tax competition is not fully warranted. One interpretation of the tax-rate bunching found for German municipalities is that some partial tax coordination among local jurisdictions is possible. Another view is that comparative performance information is relevant giving rise to yardstick competition which helps to prevent governments from extracting rents. Both interpretations, however, point to the potential for efficient policies under decentralization.

A Data Sources and Definitions

A-1 Data Description and Sources

Local Business Tax Rates: Local governments set a tax multiplier, which is then applied to a federal base rate of 3.5%. This determines the statutory business tax rate. For instance, with a multiplier of 400 percent the resulting statutory business tax rate amounts to 0.14 or 14% (= 400% × 0.035). Local multipliers are fixed for a year. The base rate of 3.5% has remained unchanged since the last tax reform in 2009. Since 2004 federal tax law demands a minimum multiplier of 200. There is no formal upper limit. The data on tax multipliers is taken from a joint publication *Hebesaetze der Realsteuern* of the federal and state statistical offices (*Statistische Aemter des Bundes und der Laender*).

Characteristics of Municipalities: Data on municipalities in Baden Wuerttemberg are obtained from the following institutions:

Regionaldatenbank Deutschland, Statistische Aemter des Bundes und der Laender
Population: population data at the end of 2011 (December 31st), structured by age.
Public debt: at the end of 2011 (December 31st), includes debt of publicly owned enterprises.

Statistisches Landesamt Baden-Wuerttemberg
Election outcomes: local elections took place on June 7th 2009. The largest total share of votes in the state (35.3%) was won by the "Freie Waehler". The second largest share was won by the Christian Democrats ("CDU") with 28.4%.
Mandates: number of seats in the municipal council.
Majority voting: some small jurisdictions have elections without party lists and use simply counts of votes for individual candidates (*Mehrheitswahl*).

Zensus 2011
Religious affiliation: the German census of 2011 provides data on religious affiliation in three categories: Roman-Catholic Church, Protestant Churches comprising Lutheran Church and Reformed Churches (*Freikirchliche Gemeinden*) and Others/None/No Information. The data refers to May 9th 2011.

Statistik der Bundesagentur fuer Arbeit
Unemployment: data on unemployment is published by the Federal Employment Agency referring to June 30th, 2011.

A-2 Variable Definitions

Urban county dummy: binary variable indicating whether a municipality is an urban county or not.

Population in absolute numbers captures the total resident population.

Demographics: old and young population refers to the population share of people under 25 and over 65 years of age.

Religious affiliation: Catholic and Protestant is the share of residents that are members of the Catholic or Protestant Churches relative to total population. The latter comprises the Lutheran Church as well as Reformed Churches.

Public debt: public debt including debt by public enterprises in Euro per 1000 inhabitants.

I(debt=0) is a binary variable indicating whether a municipality has zero public debt.

Voting system: I(majority vote) is a binary variable indicating whether a municipality has majority voting without party lists.

Mandates: number of mandates (seats) per inhabitant.

Political parties: I(*CDU*) and I(*Freie Waehler*) are binary variables indicating whether a majority in the municipal council is held by Christian Democrats (CDU) or *Freie Waehler*.

Unemployment: number of unemployed relative to all inhabitants.

Table A-1: Tax-Rate Bunching in the States of Bavaria, Brandenburg, Mecklenburg, Saxony-Anhalt

	Bavaria			Brandenburg			Mecklenburg			Saxony-Anhalt		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
I (mode)		350.4 ***	343.1 ***		137.8 ***	136.6 ***		392.8 ***	391.2 ***		37.35 ***	36.70 ***
		(14.04)	(15.74)		(2.305)	(2.433)		(3.302)	(3.493)		(1.068)	(1.105)
I (minimum)		-35.64 **	-34.43 **		2.828	3.753 *		3.828	5.598 *		-2.655 **	-2.225 **
		(14.04)	(13.82)		(2.305)	(2.257)		(3.302)	(3.118)		(1.068)	(1.090)
I (capacity)		306.4 ***	301.4 ***									
		(14.04)	(14.93)									
I (federal)		175.4 ***	169.3 ***		-1.172	-2.128		6.828 **	6.523 **		1.345	0.629
		(14.04)	(15.47)		(2.305)	(2.516)		(3.302)	(3.191)		(1.068)	(1.196)
I (div.by 10)		35.22 **	35.79 ***		5.884 **	5.940 ***		9.905 ***	9.939 ***		2.307 **	2.344 **
		(14.04)	(13.65)		(2.308)	(2.223)		(3.302)	(3.221)		(1.070)	(1.026)
m_j			-928.5 **			-136.0 *			15.22			-99.24 **
			(422.0)			(71.20)			(125.5)			(38.66)
m_j^2			438.2 **			67.14 **			0.991			47.58 ***
			(194.2)			(32.75)			(55.62)			(17.65)
m_j^3			-87.71 **			-13.91 **			-1.501			-9.629 ***
			(38.21)			(6.444)			(10.58)			(3.444)
m_j^4			6.313 **			1.028 **			0.180			0.698 ***
			(2.719)			(0.459)			(0.732)			(0.243)
Constant	6.506 ***	0.419 **	703.7 **	1.326 **	0.289 **	97.71 *	2.547 *	0.268 ***	-26.81	0.696 ***	0.349 ***	73.90 **
	(2.211)	(0.165)	(330.5)	(0.519)	(0.114)	(55.87)	(1.320)	(0.073)	(102.1)	(0.172)	(0.061)	(30.51)
R^2	0.000	0.684	0.700	0.000	0.798	0.811	0.000	0.946	0.948	0.000	0.578	0.623
R^2 corr.	0.000	0.679	0.692	0.000	0.795	0.806	0.000	0.945	0.946	0.000	0.572	0.613
Number of obs.	316	316	316	316	316	316	316	316	316	316	316	316

Dependent variable: count of municipalities that set a specific local tax multiplier. The counts refer to the empirical distribution of tax multipliers among all 2056 (Bavaria), 419 (Brandenburg), 780 (Mecklenburg), 214 (Saxony Anhalt) municipalities in the year 2011 covering a range from 200% to 515%. This results in each case in 316 observations. I (mode) characterizes the mode of the distribution of tax multipliers (340%). I (minimum) denotes the minimum tax multipliers (200%). I (capacity) identifies the reference rate used to define fiscal capacity (290%). I (federal) captures the reference rate used by the federal income tax (380%). I (div. by 10) refers to all multipliers that are multiples of 10 (200%, 210%, 220%,...). m_k captures the business tax multiplier identifying the bin. OLS regression results. Robust standard errors are given in parentheses. * denotes significance at 10%, ** significance at 5%, *** significance at 1% level.

Table A-2: Likelihood of Tax Rate Changes in the States of Bavaria, Brandenburg, Mecklenburg, Saxony-Anhalt

	Bavaria			Brandenburg			Mecklenburg			Saxony-Anhalt		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
I (mode)	-0.022	-0.022	-0.046 **	0.056	0.050	0.026	-0.224 ***	-0.205 ***	-0.270 ***	-0.162 **	-0.016	-0.099
	(0.019)	(0.020)	(0.024)	(0.034)	(0.037)	(0.051)	(0.032)	(0.035)	(0.053)	(0.075)	(0.084)	(0.101)
I (div.by 10)		0.003	0.016		0.016	0.013		-0.057	-0.017		-0.346 ***	-0.310 ***
		(0.035)	(0.036)		(0.044)	(0.046)		(0.065)	(0.070)		(0.073)	(0.080)
I (minimum)		0.000	0.000		0.001	-0.186		0.105	0.418		0.523 ***	0.270 **
		(0.000)	(0.000)		(0.108)	(0.294)		(0.102)	(0.568)		(0.054)	(0.129)
I (federal)		-0.127 ***	-0.091 ***		-0.110 ***	-0.039		0.605 ***	0.980 ***		-0.477 ***	-0.427 ***
		(0.022)	(0.033)		(0.024)	(0.050)		(0.029)	(0.060)		(0.054)	(0.089)
I (capacity)		0.030	-0.019									
		(0.022)	(0.034)									
m_i			0.309 *			-0.008			-0.065			0.107
			(0.162)			(0.020)			(0.419)			(0.175)
m_i^2			-0.001 *			0.000			0.001			-0.001
			(0.001)			(0.000)			(0.002)			(0.001)
m_i^3			0.000 *			-0.000			-0.000			0.000
			(0.000)			(0.000)			(0.000)			(0.000)
m_i^4			-0.000 *			0.000			0.000			-0.000
			(0.000)			(0.000)			(0.000)			(0.000)
Constant	0.161 ***	0.157 ***	-27.03 **	0.104 ***	0.094 **	0.982 ***	0.414 ***	0.452 ***	1.677	0.623 ***	0.823 ***	-6.667
	(0.009)	(0.033)	(13.74)	(0.019)	(0.037)	(0.046)	(0.025)	(0.058)	(33.59)	(0.040)	(0.049)	(12.53)
R^2	0.001	0.007	0.010	0.007	0.008	0.036	0.059	0.064	0.103	0.022	0.118	0.130
R^2 corr.	0.000	0.005	0.006	0.004	-0.002	0.017	0.058	0.059	0.094	0.018	0.101	0.096
Number of obs.	2056	2056	2056	417	417	417	780	780	780	214	214	214

Dependent variable: Binary variable indicating whether the tax rate set by jurisdiction i differs in 2011 from 2009. Controls refer to 2009. I (mode) is a binary variable indicating jurisdictions with tax multiplier equal to the mode of the distribution of tax multipliers (320% for Bavaria, 300% for Brandenburg and Mecklenburg, 350% for Saxony-Anhalt). I (minimum) is a binary variable indicating jurisdictions with tax multiplier equal to the minimum (200%). I (capacity) is a binary variable indicating jurisdictions with tax multiplier equal to the fixed reference rate used to define fiscal capacity (300% for Bavaria, otherwise not defined). I (federal) is a binary variable indicating jurisdictions with tax multiplier equal to the reference rate used by the federal income tax (380%). I (div. by 10) is a binary variable indicating jurisdictions with tax multiplier equal to a multiple of 10 (200%, 210%, 220%,...). m_i is the business tax multiplier set by jurisdiction i. OLS regression results. Robust standard errors in parentheses. * denotes significant at 10%, ** significant at 5%, *** significant at 1%. Observations reflect the number of municipalities. Missing values encountered due to changes in administrative structure.

References

Altshuler, R., T. Goodspeed (2014). Follow the Leader? Evidence on European and US Tax Competition. *Public Finance Review*, forthcoming.

Banerjee, A. V. (1992). A simple model of herd behavior, *Quarterly Journal of Economics*, 107(3), 797-817.

Barro, R. J. (1979). On the determination of the public debt. *Journal of Political Economy*, 87(5), 940-971.

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

Becker, S. O., P. Egger, V. Merlo (2012). How low business tax rates attract MNE activity: Municipality-level evidence from Germany, *Journal of Public Economics*, 96, 698-711.

Becker, J., R. Davis (2015), Learning to tax - interjurisdictional tax competition under incomplete information, Paper presented at the CESifo public economics conference, Munich.

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

Bordignon, M., F. Cerniglia, F. Revelli (2004). Yardstick competition in intergovernmental relationships: theory and empirical predictions, *Economics Letters*, 83, 325-333.

Brueckner, J., L. Saavedra (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 the tax rate, *Regional Science and Urban Economics*, 31, 215-245.

Buettner, T. (2003). Tax base effects and fiscal externalities of local capital taxation: evidence from a panel of German jurisdictions, *Journal of Urban Economics*, 54(1), 110-128.

Buettner, T., W. Scheffler, and A. von Schwerin (2014). Die Hebesatzpolitik bei der Gewerbesteuer nach den Unternehmensteuerreformen, *Perspektiven der Wirtschaftspolitik*, 15(4), 346-354.

Burbidge, J. B., J. DePater, G. Myers, and A. Sengupta (1997). A coalition-formation approach to equilibrium federations and trading blocs, *American Economic Review*, 87(5), 940-956.

Catenaro, M., and J. Vidal (2006). Implicit tax co-ordination under repeated policy interactions, *Louvain Economic Review*, 72(1), 5-17.

Chetty, R., J. Friedman, T. Olsen, and L. Pistaferri (2011). Adjustment Costs, Firm Responses, and Micro vs Macro Labor Supply Elasticities: Evidence from Danish Tax Records, *Quarterly Journal of Economics*, 126, 749-804.

Courchene, T. J. (1984). Equalization payments: Past, present, and future, *Ontario Economic Council*.

Devereux, M.P., B. Lockwood, M. Redoano (2008). Do countries compete over corporate tax rates?, *Journal of Public Economics*, 92(56), 1210-1235.

Dubin, R. (1995). Estimating logit models with spatial dependence, *New directions in Spatial Econometrics*, Berlin-Heidelberg, 229-242.

Egger, P., and M. Koethenbuerger (2010). Government spending and legislative organization: Quasi-experimental evidence from Germany. American Economic Journal: Applied Economics, 2(4), 200-212.

Fredriksson, P.G., D.L. Millimet (2002). Strategic interaction and the determination of environmental policy across US states, *Journal of Urban Economics*, 51(1), 101-122.

Gordon, R. H. (1992). Can capital income taxes survive in open economies?, *Journal of Finance*, 47(3), 1159-1180.

Hautsch, N., and S. Klotz (2003). Estimating the neighborhood influence on decision makers: theory and an application on the analysis of innovation decisions, *Journal of Economic Behavior and Organization*, 52(1), 97-113.

Kanbur, R., and M. Keen (1993). Jeux sans frontières: tax competition and tax coordination when countries differ in size, *American Economic Review*, 83(4), 877-892.

Keen, M., and K. Konrad (2013). The theory of international tax competition and coordination, *Handbook of Public Economics*, Volume 5, 257-328.

Keser, C., and B. Vogt (2000). Why do experimental subjects choose an equilibrium which is neither payoff nor risk dominant?, *CIRANO*.

Konrad, K. (2009). Strategy and dynamics in contests, *Oxford University Press*.

Konrad, K, and G. Schjelderup (1999). Fortress building in global tax competition, *Journal of Urban Economics*, 46(1), 156-167.

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

Musgrave, R. A. (1959). *Theory of Public Finance; a Study in Public Economy*, New York.

Olsen, A.L. (2013). The politics of digits: evidence of odd taxation, *Public Choice*, 154, 59-73.

Pope, Robin, R. Selten, and S. Kube (2009). Nominalist heuristics and economic theory, *Bonn discussion papers*, No. 17.

Razin, A., and E. Sadka (1991). International tax competition and gains from tax harmonization, *Economics Letters*, 37(1), 69-76.

Revelli, F. (2002). Local taxes, national politics and spatial interactions in English district election results, *European Journal of Political Economy* 18, 281-299.

Revelli, F. (2005). On spatial public finance empirics, *International Tax and Public Finance*, 12(4), 475-492.

Revelli, F. (2008). Performance competition in local media markets, *Journal of Public Economics*, 92 (7), 1585-1594.

Saez, E. (2010). Do taxpayers bunch at kink points?, *American Economic Journal: Economic Policy*, 2 (3),180-212.

Samuelson, P. A. (1954). The pure theory of public expenditure, *Review of economics and statistics*, 36 (4), 387-389.

Sinn, H. W. (1990). Tax harmonization and tax competition in Europe, *European Economic Review*, 34(2), 489-504.

Thomas, J., and T. Worrall (1994). Foreign direct investment and the risk of expropriation, *Review of Economic Studies*, 61(1), 81-108.

Wilson, J.D. (1986). A theory of interregional tax competition. *Journal of Urban Economics*, 19(3), 296-315.

Weingast, B. R., K. A. Shepsle, C. Johnsen (1981). The political economy of benefits and costs: A neoclassical approach to distributive politics. *Journal of Political Economy*, 89, 642-664.

Wildasin, D.E. (1987). Federal-state-local fiscal relations: a review of the treasury report, *Public Finance Quarterly*, 15 (4), 472-499.

Wildasin, D.E. (1988). Nash equilibria in models of fiscal competition, *Journal of Public Economics*, 35 (2), 229-240.

Wildasin, D. E. (1989). Interjurisdictional capital mobility: Fiscal externality and a corrective subsidy, *Journal of Urban Economics*, 25(2), 193-212.

Wildasin, D.E. (1994). Income redistribution and migration, *Canadian Journal of Economics*, 27 (3), 637-656.

Wildasin, D.E. (2003). Fiscal competition in space and time, *Journal of Public Economics*, 87(11), 2571-2588.

Vrijburg, H., and R. de Mooij (2010). Enhanced cooperation in an asymmetric model of tax competition, *CESifo Working Paper 2915.*

Wrede, M. (2001). Yardstick competition to tame the Leviathan, *European Journal of Political Economy*, 17(4), 705-721.

Zodrow, G.R., and P. Mieszkowski (1986). Pigou, Tiebout, property taxation, and the underprovision of local public goods, *Journal of Urban Economics*, 19(3), 356-370.

B A Model of Tax-Rate Signaling with Mobile Investors

We assume that capital is mobile due to firms' investment choices. Once firms have decided about investment, however, relocation costs are prohibitively high. Hence, each firm formulates an expectation regarding future taxation when deciding about investment. We consider, for simplicity, a setting with just two periods, with capital in the first period being given and capital in the second period depending on the investment decision made in the first. The demand for capital in the second period is determined by the requirement that the marginal product of capital net of taxes will have to be equal to some given rate of return

$$f'(k_{t+1}) - \mathsf{E}[\tau_{t+1}] = \rho, \tag{B-1}$$

where k_{t+1} is the capital intensity in the second period and the output per unit of labor is determined by the production function $f(k_{t+1})$. Note that there is uncertainty regarding the tax policy in the second period. For simplicity, we assume that the rate of return is known with certainty.[12] Rearranging terms, we obtain the capital-demand function

$$k_{t+1} = k(\rho + \mathsf{E}[\tau_{t+1}]).$$

Before discussing the formation of expectations by investors, let us turn to the jurisdiction's objective function. For simplicity, we do not model the local budget and how the funds are used to provide public services, and formulate the objective function using fixed preference parameters

$$\max_{\tau_t, \tau_{t+1}} \quad \Omega \quad = -\frac{1}{2}(\epsilon - \tau_t)^2 - \frac{1}{2}(\epsilon - \tau_{t+1})^2 + \delta w_{t+1} \quad \text{with } \epsilon > 0 \text{ and } \delta \geq 0. \tag{B-2}$$

[12]Following the tax competition literature (*e.g.*, Wildasin, 1988), we could replace the assumption of a given rate of return by the assumption that the total supply of capital is limited. This would allow us to endogenize the (expected) rate of return $\mathsf{E}[\rho_{t+1}]$. However, in such a general equilibrium setting, if the number of jurisdictions is sufficiently large, the effects of individual jurisdictions on the common return is small, and the results obtained below will still hold. Therefore, the analysis sticks to the simpler case with a fixed rate of return.

There are two preference parameters. ϵ is the "tax preference" of the local jurisdiction, which could reflect some fairness perception as to how much taxes should be paid by local firms. It could also reflect the demand for public services or some preference for redistribution. The quadratic terms indicate that an increasing discrepancy between the actual tax rate and the tax preference would be associated with a decline in the value of the objective function. The other preference parameter δ captures the "location preference". With $\delta > 0$ the local jurisdiction is not only interested in setting the tax rate according to its tax preference, it also cares for the earnings of labor, which is the immobile factor in our setting. Since the capital stock is given in the current period, so is the wage rate, and we can suppress the current wage rate. Hence, only the wage rate in the second period $t+1$ enters the objective function. Note that the tax preference is the same in both periods and, for simplicity, no intertemporal discount is included.

We assume that local jurisdictions have private information about their preferences. The possible types of jurisdictions and distribution of preferences are common knowledge. However, investors do not know the preferences of any specific jurisdiction and need to form beliefs. In evaluating their strategy options, local jurisdictions anticipate investors' reactions to the signal provided by the current tax policy. The Bayesian Nash-equilibrium is characterized by a set of dominant strategies for the individual jurisdictions which are mutually consistent and consistent with investors' expectation formation.

Consider first the case of jurisdictions with negligible or "weak" location preference, such that $\delta = 0$. These jurisdictions will set the tax rate equal to the tax preference in both periods. If there is variation in the tax preferences among these jurisdictions we would observe a non-degenerate distribution of tax rates. Consider, for instance, a case where ϵ is distributed symmetrically over low, medium and high tax preferences, in the sense that $\epsilon = \{\, \beta - \sigma \,,\, \beta \,,\, \beta + \sigma \,\}$ with probabilities $\{\, p \,,\, 1 - 2p \,,\, p \,\}$, $0 < p < 0.5$ respectively. From the objective function (B-2) it is straightforward to show that jurisdictions with weak location preference $\delta = 0$ will have low, medium, or high taxes, according to their tax preference.

For jurisdictions with a non-trivial or "strong" location preference $\delta > 0$, however, the optimal tax policy depends on the effect of the tax policy on location. Formally, in period $t+1$, the wage rate is determined by the marginal product

$$w_{t+1} = f\left(k_{t+1}\right) - f'\left(k_{t+1}\right) k_{t+1}.$$

Since period $t+1$ is the last period, and since location decisions have already been made, in t+1 the local jurisdiction chooses the tax rate equal to the preference parameter

$$\tau_{t+1} - \epsilon \stackrel{!}{=} 0.$$

However, the tax policy chosen in period t may have an effect on the tax rate expected for period $t+1$. Formally, in period t the first-order condition for the optimal tax rate is

$$\left(\epsilon - \tau_t\right) + \delta w'\left(k_{t+1}\right) \frac{\partial k_{t+1}}{\partial E\left[\tau_{t+1}\right]} \frac{dE\left[\tau_{t+1}\right]}{d\tau_t} \stackrel{!}{=} 0.$$

Since $w'\left(k_{t+1}\right) = -f''\left(k_{t+1}\right) k_{t+1}$ and $\frac{\partial k_{t+1}}{\partial E[\tau_{t+1}]} = \frac{1}{f''(k_{t+1})}$ the optimal policy is characterized by

$$\left(\epsilon - \tau_t\right) - \delta k_{t+1} \frac{dE\left[\tau_{t+1}\right]}{d\tau_t} \stackrel{!}{=} 0.$$

Obviously, whenever the expected tax rate is positively related with the choice of the current tax rate $\frac{dE[\tau_{t+1}]}{d\tau_t} > 0$, and if $\delta > 0$, the jurisdiction may deviate from its tax preference by choosing a tax rate $\tau_t < \epsilon$. The question is, however, under which conditions investor expectations are affected by tax policy signals in period t.

The expectation formation of investors will crucially depend on the preference distribution of jurisdictions. If there are plenty of jurisdictions with strong location preference, a certain tax rate might not be a very informative signal, as it does not reveal the tax preference. Instead, if most jurisdictions have a weak location preference, setting a low tax rate in period t may substantially lower the expected tax rate and hence exert positive effects on investment. But also the joint distribution of tax and location preferences matters. If jurisdictions with a strong location preference have a

Table B-1: Assumed Preference Distribution

			Location preference		
			Weak ($\delta = 0$)	Strong ($\delta > 0$)	All
Tax preference	Low	($\epsilon = \beta - \sigma$)	$p\,(1 - \pi)$	0	$p\,(1 - \pi)$
	Medium	($\epsilon = \beta$)	$(1 - 2p)\,(1 - \pi)$	0	$(1 - 2p)\,(1 - \pi)$
	High	($\epsilon = \beta + \sigma$)	$p\,(1 - \pi)$	π	$\pi + p\,(1 - \pi)$
	All		$1 - \pi$	π	1

low tax preference, just sticking to the low tax rate would be the best choice anyway. However, if jurisdictions with strong location preference happen to have a medium or high tax preference, the trade-off between meeting the tax preference right in the first period and sending out a signal to investors that results in higher future capital will be important. To simplify matters, let us focus on the latter case and assume the specific distribution of preferences depicted in table B-1.

With this preference distribution, a share π of jurisdictions has a strong location preference with $\delta > 0$ as well as a high tax preference $\epsilon = \beta + \sigma$. With regard to the other jurisdictions we have a symmetric distribution of tax preferences around β. In order to make sure that all three tax preferences are represented in the resulting tax rate distribution we need to restrict $0 < p < 0.5$. Note that this includes as a possibility a uniform distribution of tax preferences $p = 1/3$ among jurisdictions with weak location preferences.

B-1 Expectation Formation when Location Preference is Known

As we have seen above, jurisdictions with weak location preference $\delta = 0$ will always stick to their tax preference. Hence, forming expectations about tax rates is straightforward. If the investor would know that a jurisdiction has no strong location preference, he/she just expects the same tax rate to prevail in the future. Formally, we have the following conditional expectations for the tax rate in period $t+1$:

$$\begin{aligned} \mathsf{E}\left[\tau_{t+1} | \tau_t = \beta - \sigma \wedge \delta = 0\right] &= \beta - \sigma \\ \mathsf{E}\left[\tau_{t+1} | \tau_t = \beta \wedge \delta = 0\right] &= \beta \\ \mathsf{E}\left[\tau_{t+1} | \tau_t = \beta + \sigma \wedge \delta = 0\right] &= \beta + \sigma. \end{aligned}$$

If the investor would know that the jurisdiction has a strong location preference, given our assumption about the preference distribution, he/she will expect to observe a high tax rate in the future, regardless of the current tax rate

$$\mathsf{E}\left[\tau_{t+1} | \delta > 0\right] = \beta + \sigma.$$

With imperfect information, it is necessary to explore how the information on current tax policy affects the investor beliefs.

B-2 Expectation Formation with Unknown Location Preference

Investors know that jurisdictions with strong location preference $\delta > 0$ have three options.

- S1: The first option is to set the tax rate in period t equal to the tax preference $\tau_t = \beta + \sigma$.
- S2: Jurisdictions could choose a tax rate that mimics the tax policy of medium tax preference jurisdictions, and set $\tau_t = \beta$.
- S3: A third option is to mimic the tax policy of low tax preference jurisdictions, and set $\tau_t = \beta - \sigma$.

One might wonder, whether jurisdictions with strong location preference might choose tax rates other than $\beta - \sigma$, β or $\beta + \sigma$. Since other tax rates are never chosen by jurisdictions with weak location preference, such a choice would immediately reveal the type of the jurisdiction, and is therefore strictly dominated by strategy S1.

Lemma: In a setting with heterogeneous jurisdictions with objective function (B-2), with tax and location preferences distributed according to table B-1 and if the distribution of preferences is common knowledge, investors with incomplete information about the type of a jurisdiction will base their expectations about future tax rates on current choices. Even if investors build their prediction on the assumption that jurisdictions with strong location preference all deviate from their tax preference, the tax rate expected by investors is reduced under this strategy.

Proof: If investors assume that jurisdictions with strong location preference follow strategy $\mathsf{S1}$, the investor's conditional expectations given the observed tax rate and the belief about the strategy is straightforward

$$\begin{aligned} \mathsf{E}\left[\tau_{t+1} \middle| \tau_t = \beta - \sigma, \mathsf{S1}\right] &= \beta - \sigma \\ \mathsf{E}\left[\tau_{t+1} \middle| \tau_t = \beta, \mathsf{S1}\right] &= \beta \\ \mathsf{E}\left[\tau_{t+1} \middle| \tau_t = \beta + \sigma, \mathsf{S1}\right] &= \beta + \sigma. \end{aligned} \tag{B-3}$$

Obviously, under the investor belief, setting the tax rate to β in period t clearly lowers the expected tax rate relative to a choice of $\tau_t = \beta + \sigma$.

Suppose investors assume that jurisdictions with strong location preference follow strategy $\mathsf{S2}$ which involves setting the tax rate equal to the expected value of the tax preference distribution of jurisdictions with weak location preference. In this case, the conditional investor expectations are

$$\begin{aligned} \mathsf{E}\left[\tau_{t+1} \middle| \tau_t = \beta - \sigma, \mathsf{S2}\right] &= \beta - \sigma \\ \mathsf{E}\left[\tau_{t+1} \middle| \tau_t = \beta, \mathsf{S2}\right] &= \beta + \sigma \left(\frac{\pi}{(1-2p)(1-\pi) + \pi} \right) \\ \mathsf{E}\left[\tau_{t+1} \middle| \tau_t = \beta + \sigma, \mathsf{S2}\right] &= \beta + \sigma. \end{aligned} \tag{B-4}$$

Since $p < 0.5$, we know that setting the tax rate equal to β will contribute to lower tax expectations than a choice of $\tau_t = \beta + \sigma$, if the share of jurisdictions with strong location preference π is below unity.

Investors may also assume that jurisdictions with strong location preference follow strategy S3, and imitate the tax policy of a jurisdiction at the lower end of the tax preference distribution of jurisdictions with weak location preference. In this case, the conditional investor expectations are

$$\begin{aligned} \mathsf{E}\left[\tau_{t+1}\middle|\,\tau_t=\beta-\sigma,\mathsf{S3}\right] &= \beta-\sigma+2\sigma\left(\frac{\pi}{p\left(1-\pi\right)+\pi}\right) \\ \mathsf{E}\left[\tau_{t+1}\middle|\,\tau_t=\beta,\mathsf{S3}\right] &= \beta \\ \mathsf{E}\left[\tau_{t+1}\middle|\,\tau_t=\beta+\sigma,\mathsf{S3}\right] &= \beta+\sigma. \end{aligned} \tag{B-5}$$

Since $p > 0$, setting the tax rate equal to the lower end of the tax preference distribution $\tau_t = \beta - \sigma$ will also induce a decline in tax expectations relative to a choice $\tau_t = \beta + \sigma$. ∎

B-3 Policy Options and Strategic Choice

As we have seen above, a jurisdiction with strong location preference can choose among different options. It may just follow its original preference or it might strategically choose a tax rate according to the distribution of tax preferences, as it is concerned about future wages ($\delta > 0$). We have shown that a deviation from the preferred tax rate will induce lower expectations even though investors take into account that tax policy could be reversed in the future. However, whether such a strategic policy is ultimately beneficial depends on costs and benefits. More specifically, it depends on the objective function (see equation (B-2)), whether a jurisdiction gains by deviating from its tax preference in the first period, and if so, which strategy is chosen. To show that a strategic policy may indeed be associated with the highest value of the objective function, we insert the first-period choice of the tax rate as well as the implied expected tax rate into the objective function to obtain an indirect objective function.

To simplify matters we consider a quadratic production function $f\left(k_t\right) = k_t - \frac{\alpha}{2}k_t^2$ with $f'\left(k_t\right) = 1 - \alpha k_t$. With this function

$$w\left(k_{t+1}\right) = \frac{\alpha}{2}k_{t+1}^2 = \frac{\alpha}{2}\left(\frac{1}{\alpha}\left(1 - E\left[\tau_{t+1}\middle|\,\tau_t\left(s\right), S\right] - \rho\right)\right)^2 = \frac{1}{2\alpha}\left(1 - E\left[\tau_{t+1}\middle|\,\tau_t\left(s\right), S\right] - \rho\right)^2,$$

where s denotes the choice of the jurisdiction, and S the investor belief, $s, S \in \{\mathsf{S1}\ ;\ \mathsf{S2}\ ;\ \mathsf{S3}\}$, and

$$\tau_t(s) = \begin{cases} \tau_t(\mathsf{S1}) & = & \beta + \sigma \\ \tau_t(\mathsf{S2}) & = & \beta \\ \tau_t(\mathsf{S3}) & = & \beta - \sigma. \end{cases}$$

Since the tax rate in the second period is always chosen in accordance with the tax preference $\tau_{t+1} = \beta + \sigma$, it does not contribute to the indirect utility function. Hence, the indirect utility of choosing strategy s conditional on the investor belief that the strong location preference jurisdiction follows strategy S is equal to

$$\Omega(s, S) = -\frac{1}{2}(\beta + \sigma - \tau_t(s))^2 + \frac{\delta}{2\alpha}(1 - E[\tau_{t+1} | \tau_t(s), S] - \rho)^2.$$

Now, we are in a position to evaluate the different options.

Proposition: In a setting with heterogeneous jurisdictions with objective function (B-2), with asymmetric information, preferences distributed according to table B-1, with quadratic production functions, and investor beliefs as specified in equations (B-3,B-4,B-5), the best strategy of a high tax jurisdiction with strong location preference may be to deviate from the preferred tax rate in the first period. With $0 < p < 0.5$, if the location preference δ is within a certain range and if the number of jurisdictions with strong location preference is limited, the Bayesian Nash-equilibrium is a pooling equilibrium, where jurisdictions with strong location preference set their first-period tax rate equal to the median of the tax rate distribution.

Proof: To proof that this pooling equilibrium exists, we show first that the utility of any jurisdiction with strong location preference is higher even if all these jurisdictions pursue strategy S2 rather than S1 and if investors rightly predict this strategy. Secondly, we show that the utility with this strategy is also higher relative to strategy S3 when investors make right predictions. In a third step, we use the above formulation of beliefs to show that strategy S2 strictly dominates all other choices such that other equilibria can be ruled out.

1. $\Omega(\mathsf{S2},\mathsf{S2}) > \Omega(\mathsf{S1},\mathsf{S1})$ Consider the case where jurisdictions with strong location preference sets their tax rate according to the tax preference as anticipated by investors $s = \mathsf{S1}, S = \mathsf{S1}$. In this case, the tax rate in t would be $\tau_t = \beta + \sigma$ and as shown above, the expected tax rate would be the same. The corresponding value of the indirect objective function for an individual jurisdiction is

$$\Omega(\mathsf{S1},\mathsf{S1}) = \frac{\delta}{2\alpha}(1-\beta-\sigma-\rho)^2 .$$

If jurisdictions with strong location preference follow the second strategy and the investors rightly anticipate this $s = \mathsf{S2}, S = \mathsf{S2}$, the tax rate in t would be $\tau_t = \beta$ and, as shown above, the expected rate would be $E[\tau_{t+1}|\,\tau_t = \beta, \mathsf{S2}] = \beta + \sigma\left(\frac{\pi}{(1-\pi)(1-2p)+\pi}\right)$. In this case the objective function has the value

$$\Omega(\mathsf{S2},\mathsf{S2}) = -\frac{1}{2}\sigma^2 + \frac{\delta}{2\alpha}\left(1-\beta-\sigma\left(\frac{\pi}{(1-\pi)(1-2p)+\pi}\right)-\rho\right)^2 .$$

It is now straightforward to show that

$$\Omega(\mathsf{S2},\mathsf{S2}) > \Omega(\mathsf{S1},\mathsf{S1}) \tag{B-6}$$

holds if δ is sufficiently large, formally

$$\delta > \underline{\delta} \equiv \left[\frac{\alpha}{\frac{2\kappa}{\sigma}(1-\varphi)-(1-\varphi^2)}\right], \text{ where } \kappa \equiv 1-\beta-\rho \text{ and } \varphi \equiv \frac{\pi}{(1-\pi)(1-2p)+\pi}. \tag{B-7}$$

Note that positive investment requires $\alpha > 0$ and $\kappa > \sigma$.

2. $\Omega(\mathsf{S2},\mathsf{S2}) > \Omega(\mathsf{S3},\mathsf{S3})$ To consider the utility under strategy $\mathsf{S3}$, note first that the utility obtained under this strategy would be higher if investors wrongly believe that the jurisdictions with strong location preference follow strategy $\mathsf{S2}$ rather than $\mathsf{S3}$, $\Omega(\mathsf{S3},\mathsf{S2}) > \Omega(\mathsf{S3},\mathsf{S3})$, as the expected tax rate would be lower. Hence to show that $\Omega(\mathsf{S2},\mathsf{S2})$ dominates $\Omega(\mathsf{S3},\mathsf{S3})$, it suffices to show that $\Omega(\mathsf{S2},\mathsf{S2}) > \Omega(\mathsf{S3},\mathsf{S2})$. In the latter case, the corresponding value of the indirect objective function is

$$\Omega(\mathsf{S3},\mathsf{S2}) = -\frac{1}{2}(2\sigma)^2 + \frac{\delta}{2\alpha}(1-\beta+\sigma-\rho)^2 .$$

Now it is straightforward to show that $\Omega(\mathsf{S2},\mathsf{S2}) > \Omega(\mathsf{S3},\mathsf{S2})$ holds, if δ is sufficiently small. Hence, with

$$\delta < \overline{\delta} = \left[\frac{3\alpha}{\frac{2\kappa}{\sigma}(1+\varphi)+(1-\varphi^2)}\right], \tag{B-8}$$

we know that

$$\Omega(\mathsf{S2},\mathsf{S2}) > \Omega(\mathsf{S3},\mathsf{S3}). \tag{B-9}$$

Strict dominance of $\mathsf{S2}$ Now we are prepared to show that in a setting with $\underline{\delta} < \delta < \overline{\delta}$, strategy $\mathsf{S2}$ clearly dominates the other strategies.[13] To see this, note that the gain from playing strategy $\mathsf{S2}$ is highest if investors would wrongly believe that jurisdictions with strong location preference play $\mathsf{S1}$. Therefore, $\Omega(\mathsf{S2},\mathsf{S1}) > \Omega(\mathsf{S2},\mathsf{S2})$ and using equation (B-6) we have

$$\Omega(\mathsf{S2},\mathsf{S1}) > \Omega(\mathsf{S1},\mathsf{S1}). \tag{B-10}$$

Under the investor belief $\mathsf{S1}$, playing strategy $\mathsf{S2}$ also yields a higher benefit than strategy $\mathsf{S3}$. Since

$$\Omega(\mathsf{S2},\mathsf{S1}) > \Omega(\mathsf{S3},\mathsf{S1}) \tag{B-11}$$

requires

$$-\frac{1}{2}\sigma^2 + \frac{\delta}{2\alpha}(1-\beta-\rho)^2 > -\frac{1}{2}(2\sigma)^2 + \frac{\delta}{2\alpha}(1-\beta+\sigma-\rho)^2.$$

Reformulation yields

$$\delta < \frac{3\alpha}{2\frac{\kappa}{\sigma}+1}.$$

Taking account of condition (B-8) this holds for small values of φ.

Dominance also holds if investors believe that jurisdictions with strong location preference follow

[13] To ensure that the interval for δ is not empty, φ needs to be small. With $0 < p < 0.5$, this requires π to be sufficiently small.

strategy $\mathsf{S2}$. To see this, note first that we have already shown that

$$\Omega\left(\mathsf{S2},\mathsf{S2}\right) > \Omega\left(\mathsf{S3},\mathsf{S2}\right) \tag{B-12}$$

under the restriction for δ. Choosing $\mathsf{S1}$ is also not preferred. As investors would expect a high tax rate, $\Omega\left(\mathsf{S1},\mathsf{S2}\right) = \Omega\left(\mathsf{S1},\mathsf{S1}\right)$. From above, $\Omega\left(\mathsf{S2},\mathsf{S2}\right) > \Omega\left(\mathsf{S1},\mathsf{S1}\right)$, hence

$$\Omega\left(\mathsf{S2},\mathsf{S2}\right) > \Omega\left(\mathsf{S1},\mathsf{S2}\right). \tag{B-13}$$

If investors believe that jurisdictions with strong location preference follow strategy $\mathsf{S3}$, strategy $\mathsf{S2}$ still dominates under the restrictions for δ. To see this note that $\Omega\left(\mathsf{S2},\mathsf{S3}\right) > \Omega\left(\mathsf{S2},\mathsf{S2}\right)$. The latter is preferred to $\Omega\left(\mathsf{S3},\mathsf{S3}\right)$ and, thus,

$$\Omega\left(\mathsf{S2},\mathsf{S3}\right) > \Omega\left(\mathsf{S3},\mathsf{S3}\right). \tag{B-14}$$

Again, $\mathsf{S1}$ is no option. Since investors would expect a high tax rate, $\Omega\left(\mathsf{S1},\mathsf{S3}\right) = \Omega\left(\mathsf{S1},\mathsf{S1}\right)$. From above, $\Omega\left(\mathsf{S2},\mathsf{S2}\right) > \Omega\left(\mathsf{S1},\mathsf{S1}\right)$, and $\Omega\left(\mathsf{S2},\mathsf{S3}\right) > \Omega\left(\mathsf{S2},\mathsf{S2}\right)$, and, hence,

$$\Omega\left(\mathsf{S2},\mathsf{S3}\right) > \Omega\left(\mathsf{S1},\mathsf{S3}\right). \tag{B-15}$$

Given the six inequalities (B-10) - (B-15), $\mathsf{S2}$ strictly dominates the other strategies regardless of whether or not investor beliefs are consistent with the equilibrium path.

■