Multilevel taxation, competition, and sorting: Evidence from regional borders

Federico Revelli[*]   Roberto Zotti[†]

Abstract

This paper exploits the multi-level structure of progressive personal income taxation in Italy, and particularly the discontinuity of upper-tier top marginal rates at regional borders, to estimate the response of the share of high income taxpayers to local tax differentials and to analyze the tax rate setting decisions of lower-tier municipal authorities relative to their neighbors. We discuss the properties and data requirements of recently proposed border-discontinuity instrumental variable estimators of spatial reaction functions that use cross-border upper-tier tax policies as instruments for lower-tier ones (Parchet R., "Are local tax rates strategic complements or strategic substitutes?" *American Economic Journal: Economic Policy* (2019) forthcoming). Using a large panel dataset of municipal and regional income tax rates spanning through two decades, we find a positive and significant impact of the net-of-tax rate on the share of top income taxpayers that reside in a locality. On the other hand, the evidence of negative spatial correlation in local income tax rates turns out to be highly sensitive to the specification of the empirical model, leaving open the issue of whether they are strategic complements or substitutes.

JEL classification: H24; H71; R23
Key words: local income taxation; fiscal competition; sorting.

[*]Department of Economics and Statistics "*S. Cognetti de Martiis*," University of Torino, Campus Luigi Einaudi, Lungo Dora Siena 100A, 10153 Turin (Italy); e-mail: federico.revelli@unito.it.

[†]Department of Economics and Statistics "*S. Cognetti de Martiis*," University of Torino, Campus Luigi Einaudi, Lungo Dora Siena 100A, 10153 Turin (Italy); e-mail: roberto.zotti@unito.it.

1 Introduction

In the absence of formal or substantive barriers to internal migration, decentralized taxation of personal income - a common feature of fiscal structures both in North America and in Europe - can exert a decisive influence on the location of households, particularly those at the upper end of the income distribution. From an empirical point of view, two major econometric issues need to be tackled when studying the impact of income taxation on the location of households. The first is that of reverse causality: the tax income schedule that a locality can afford to set depends itself on the income distribution of the population that has chosen to reside in that jurisdiction. Second, in the presence of income tax base mobility local tax setting decisions will end up being interdependent due to the fiscal externality that each taxing authority inflicts on the other authorities, most likely the ones that are geographically close. As a result, a growing recent literature attempts at identifying the impact of local tax structures on individual location decisions by exploiting exogenous shocks to local income taxation schedules. These shocks can arise either from the presence of higher-level administrative units that create discontinuities at internal borders, from the existence of legal thresholds in the computation of income tax liabilities, or from spatially-limited reforms creating geographically defined control and treatment groups.

This paper uses the multi-level structure of income taxation in Italy, and particularly the discontinuity of top marginal income tax rates of progressive schedules at regional borders, to analyze the response of top income taxpayers to local tax differentials and the income tax rate setting decisions of lower-tier authorities (municipalities). As for the first issue, we employ the upper-tier regional income tax rate as an instrument for the consolidated municipal-regional tax rate, and find a positive and significant impact of the net-of-tax rate on the share of top income taxpayers. Next, we apply the border-discontinuity instrumental variable estimator proposed by Parchet (2019) and discuss its properties and implications by comparing it to conventional estimators of spatial reaction functions. The estimator is applicable to data on localities whose set of first-order geographic neighbors includes 'foreign' localities, that is units that are located in a different region, and proposes to use the tax policy of the 'foreign' regional authority as an instrument for the potentially endogenous weighted average of lower-tier neighbors' fiscal policies - the conventional spa-

tial lag in a horizontal spatial reaction function. Using a large panel dataset of municipal and regional income tax rates spanning through two decades, we find that border-discontinuity instrumental variable estimators that use cross-border upper-tier tax policies as instruments for lower-tier ones return evidence of negative spatial correlation that is highly sensitive to the specification of the empirical model, leaving open the issue of whether local income tax rates are strategic complements or substitutes.

The rest of the paper is structured as follows. Section **2** reviews the recent literature on local income tax competition for mobile taxpayers. Section **3** develops the empirical model and outlines our estimation strategy. Section **4** describes the Italian institutional context and the dataset, section **5** presents the results, and section **6** concludes.

2 Local income taxation, competition, and sorting: a review

It has been common in the past couple of decades to address the issue of whether decentralized authorities interact strategically with each other in fiscal policy-making by estimating spatial tax reaction functions. A spatial tax reaction function allows the observations on a tax policy instrument at a number of relevant spatial locations (states in federal countries, regions, or municipalities) to be determined simultaneously as the vector of tax policy realizations in contiguous or related localities through a first or higher order spatial auto-regressive specification that is assumed to accurately describe the underlying data generating process. In reality, ascertaining whether the geographical location of local authorities plays a role in their tax setting decisions has proved extremely challenging for a number of reasons (Revelli, 2005). First, in the presence of geographically-coded observations that might not be independent of one another, characterizing and interpreting the dependence among those observations necessarily requires imposing restrictions on the process linking the observations at various locations. Given that estimating the full variance-covariance structure is not feasible, the conventional spatial econometrics approach consists in imposing a priori a spatial weights matrix postulating the links – or the absence of links – between any two pairs of observations, and estimate the extent of spatial auto-correlation among units conditional on that specific spatial structure (Gibbons and Overman, 2012). Second, estimating the effect that

3

the realization of a given variable has on the realizations of that variable in related locations requires addressing the inherent simultaneity issue. Finally, one can think of a number of ways that events occurring at nearby locations might not be independent. These include an endogenous interaction effect, by which the behavior of a policy-maker at a given location affects the behavior of other policy-makers in related locations, or an exogenous effect arising from the impact of the average underlying traits of the 'neighborhood' on individual outcomes, or to 'correlated effects' due to unobserved characteristics driving agents' behavior (Manski, 1993). Identification of the parameters capturing those effects will tend to be difficult in most circumstances.

As a result, credible identification of the causal impact of the spatial lag of an endogenous variable on own decisions is increasingly obtained by exploiting exogenous shocks to neighbors' policies. These can arise from changes in rules, limits, and resource distribution mechanisms having an impact on neighbors' decisions, while having no direct effect on own decisions. Lyytikainen (2012) represents one of the first attempts in this direction. He makes use of a local public finance reform that occurred in Finland in 2000 and raised the upper and lower limits that the central government imposed on local governments' property tax rates. For all authorities that were setting in 1999 a lower tax rate than the new limit, the reform generated exogenous and large tax rate changes given that the new limit was almost twice as large as the previous one. Those forced increases can be taken as the required source of exogenous variation and can be used as instruments for changes in neighbors' tax rates in a spatial lag specification of a property tax rate determination equation.[1] The empirical analysis conducted on the differenced cross-section 1999-2000 reveals virtually no response of unconstrained authorities to tax rate changes taking place in the neighborhood. These results stand in contrast with those that are obtained when estimating the spatial reaction function either by spatial maximum likelihood methods or by an instrumental variables approach that uses neighbors' socio-demographic attributes as instruments for their tax rates. Both of those approaches yield large and significant positive estimates of the

---

[1]In the empirical implementation of his approach, Lyytikainen (2012) realizes that corner solutions arise when authorities are against the new lower limit in the year 2000, thus complicating the estimation of the spatial reaction function (Di Porto and Revelli, 2013). Consequently, he focuses on the behavior of municipalities for which the new lower limit is not binding (i.e., drops the corner solution observations), and estimates the response of those unconstrained authorities to the tax rate changes in all neighboring authorities, whether or not they are constrained by the new limit, using as instruments the predicted imposed tax rate increases corresponding to corner solution observations.

first-order spatial auto-correlation coefficient capturing the slope of the spatial reaction function, suggesting that they might overestimate the degree of spatial interdependence.

Following a similar logic, Isen (2014) exploits the widespread use of referenda by local governments imposed by state law in Ohio for the determination of taxation and public spending to overcome the problem of simultaneous determination of own and neighboring authorities' fiscal decisions in the fiscal reaction function. In particular, Isen (2014) adopts a regression discontinuity approach, in the sense that he exploits the outcomes of 'close' referenda, where one can assume a quasi-random assignment of 'treatment' to fiscal decisions of neighbors that just passed the approval threshold relative to the measures that just narrowly failed to pass. Using purely geographic as well as migration-based and socio-demographic similarity criteria of proximity between local jurisdictions (counties, municipalities, school districts) and employing a number of distinct dependent variables (rates, revenues, expenditures), Isen (2014) finds no evidence of fiscal interaction. Interestingly, when the same data are used to estimate a fiscal reaction function by means of an instrumental variables approach using neighbors' socio-demographic characteristics as instruments for neighbors' fiscal policies, the estimates of the spatial interaction effects turn out to be large and statistically significant.

Baskaran (2014) adopts a similar approach in order to consistently estimate a reaction function for local government fiscal policies. He analyses the impact of a reform in a German state, using a nearby state as control. In particular, he exploits the exogenous reform of the local fiscal equalization scheme in the German state of North Rhine-Westphalia to study the mimicking behavior by the authorities in the neighboring state of Lower Saxony, that were not directly affected by the reform. The exogenous shock to the former authorities, that were induced to raise their tax rates by the reform, is used to identify the presence of fiscal interactions by overcoming the issues of simultaneous determination and endogeneity that typically arise in studies of this type. Baskaran (2014) finds no evidence of tax mimicking when using the exogenous variation caused by the reform as instruments, irrespective of the weighting scheme used. On the other hand, he finds too that estimation of a tax reaction function that uses neighboring authorities' demographic and political characteristics as instruments for their tax policies returns evidence of strategic interaction that might in reality be caused by the use of invalid instruments.

Finally, Parchet (2019) tackles the issue of the sign of an income tax reaction function in a multi-tiered structure of government. First, his theoretical model shows that the slope of the local tax reaction function can either be positive (tax rates of competing jurisdictions are strategic complements) or negative (strategic substitutes), and highlights the factors that determine the sign in terms of economies of scale in the production of the public good and the elasticity of the marginal utility of the public good. He then exploits the discontinuity at internal Swiss borders to estimate a local income tax reaction function, where neighboring localities' income tax rates across the border are instrumented with the upper level of government (cantonal) tax rate. The idea is that fiscal reforms at the cantonal level provide a one-way and arguably exogenous source of variation for the tax rate set by the local level of government, in the sense that the local fiscal policy does not in turn affect the state-level one. In addition, taxpayers close to the state border are mobile, and make their decisions in response to their total income tax rate. He estimates a linear spatial reaction function where the consolidated local-state income tax rate measured at the municipal level is allowed to depend on the weighted average of the consolidated local-state income tax rates of neighboring municipalities, that are possibly located in different cantons. Panel data IV estimation using neighbors' characteristics as instruments based on a subset of cantonal-border municipalities (localities within a road distance of 10km from a neighbor in a different canton) and the consolidated tax rates on top income taxpayers returns a spatial auto-correlation coefficient of almost 0.7, well in excess of the OLS estimate. On the other hand, when using neighboring cantonal tax rates multiplied by the share of municipalities located in those cantons as instruments for average neighbors' tax rates, the key interaction coefficient turns negative (-0.5). In fact, weighting the native instrument (the neighboring cantons' tax rates) by the shares of municipalities that are respectively located in those cantons turns out to be crucial for the relevance of the first-stage regression, a point we come back to in the next section.

Indeed, interdependence in local income tax policy decisions among decentralized authorities rests on the hypothesis of mobility of households in response to fiscal differentials. From an empirical point of view, though, addressing the issue of whether local fiscal policies play a significant role in households' internal migration decisions besides the well-established 'push and pull' factors from labor market conditions and environmental amenities has proved problematic

6

too, and thwarted by endogeneity, reverse causality, and measurement issues (Kuminoff et al., 2013). One early way of overcoming the policy endogeneity problem has been to use household level data and employ microeconometric techniques to identify the variables affecting individual decisions of where to reside. The argument is that, from the perspective of an individual household, the characteristics of the potential destinations can be taken as exogenously given. However, and in spite of being applied to the same institutional environment (the multi-tiered financial structure of the Swiss federation, the OECD country with the largest cross-community variation in local income tax rates), the early papers using this approach did not come to univocal conclusions on the relevance of local tax differentials in household sorting processes. Schmidheiny (2006) used data from the metropolitan area of Basel and, employing a random utility maximization empirical model that exploited the progressivity of the local income tax schedules, found that, besides social interaction and commuting time motives, taxes played an important role in location decisions, with rich households being substantially more likely to move to low-tax jurisdictions than poor ones. On the other hand, based on micro-data from the Swiss Household Panel covering the whole of Switzerland, Liebig and Sousa-Poza (2006) found no evidence of tax-induced internal migration, but rather found that internal migration decisions were mostly influenced by housing market and accommodation-related factors. Based on the 2000 Swiss Census, Liebig et al. (2007) studied the migration decisions of distinct socio-demographic groups as a function of communities' tax rates, expenditures on various public services (including education, health, transport, and the environment), local amenities, and housing prices, finding a small effect of taxes, and only on young and highly educated individuals. Finally, Schaltegger et al. (2011) used the shares of four income classes in the local communities within the Swiss canton of Zurich (from below 30,000 to over 90,000 Swiss francs) and regressed them on communities' tax multipliers, average tax multipliers in neighboring communities, and land prices, using a vector of geographic and political characteristics as instruments. They found a significant negative (positive) impact of local tax multipliers on the share of the highest (lowest) income group, and opposite-signed effects from land prices and from the average tax multipliers in the neighborhood.[2]

_____

[2] As far as the US are concerned, Young and Warner (2011) and Young et al. (2016) study the outmigration impact of the imposition of state-level "millionaire taxes" on top income-earners. Their results generally suggest little responsiveness of the tax base, with estimated elasticities that are small and only marginally significant either when analyzing a specific

More recently, a growing literature attempts at identifying the impact of the local tax structure on individual location decisions by exploiting exogenous shocks to local income taxation schedules. Martinez (2017) analyses the impact of a regressive income tax reform in the canton of Obwalden (central Switzerland) in 2006 on taxpayers' location decisions, using individual income tax data and exact moving dates. The reform, approved by popular referendum, changed the existing flat income tax schedule into a regressive tax scheme (declining marginal income tax rates for a range of high incomes) that mostly benefitted taxpayers in the top 1% of the income distribution in the first two years. A Federal Court rule later imposed the return to a flat income tax rate, though the Canton chose a lower level than the pre-existing one, in 2008. Martinez (2017) shows by a difference-in-differences approach using nearby cantons as controls that the share of high-income taxpayers in the Canton and the level of taxable income per taxpayer both raised after the reform. When investigating the mechanism leading to those results, she finds that the effect is mostly due to rich taxpayers moving in, a phenomenon facilitated by the lack of cultural barriers, the small distances between cantons, and the absence of legal restrictions. From the point of view of cantonal revenues, though, the reform was roughly revenue neutral.

Basten et al. (2017) estimate the elasticity of housing rents to local income tax differentials by means of a discrete choice model where households sort into localities according to their incomes, tastes for housing, taxes, and neighborhood characteristics. Based on detailed information on all residences advertised for rent in Switzerland over the years 2005-2012 as well as on individual socio-demographics characteristics and taxable income data, they adopt a boundary discontinuity design. This implies comparing close-by residences sharing the same amenities and neighborhood characteristics, but facing different local income taxes because of being located on either side of an inter-municipal border. First, they find that the rent elasticity to the local income tax rate is about

---

New Jersey tax that raised the tax rate on top incomes by 2.6 percentage points, or when looking at all million-dollar income earners in the US for more than a decade. Moretti and Wilson (2017) use data on US state personal income tax, corporate income tax, R&D tax credits and investment tax credits to estimate to what extent the geographical allocation of highly skilled workers (star scientists) is affected by the large time-series and cross-sectional variation in fiscal burdens. In particular, to control for longstanding state characteristics that can produce spurious correlation between tax rates and number of highly skilled workers, they estimate the elasticity of migration to fiscal-related incentives by regressing changes in the number of scientists moving between states on the inter-state tax differentials. They find significant and large elasticities of the flow of scientists with respect to personal as well as corporate taxes and investment tax credits.

-0.3, corresponding to two thirds of house price capitalization estimates from conventional hedonic regression models that might not fully account for unobserved location and neighborhood characteristics. Second, their evidence points to household spatial sorting according to income. High-income households are willing to pay higher rents than poorer ones for residing in neighborhoods that are wealthier, that have lower income taxes, and that require lower moving distance, consistent with the hypothesis of non-homothetic preferences.

Schmidheiny and Slotwinski (2018) study the tax-induced behavior of high-income foreigners in Switzerland. They exploit a special regulation in the Swiss tax law according to which foreigners with gross income below a threshold pay a fixed cantonal rate for the first five years of stay, while those above the threshold are immediately subject to the ordinary regime, that has variation in tax rates (local multipliers applied on cantonal rates) across municipalities. The existence of the income threshold and of the duration threshold makes it possible to employ regression discontinuity designs. First, they hypothesize that newly arrived foreigners with income below the threshold should choose to locate in high tax municipalities, because they can enjoy the lower housing prices while not being subject to the high income tax rates because of the special fiscal regime. On the other hand, foreigners with income above the threshold should choose to locate in low tax municipalities to minimize their tax burden. In addition, one should observe taxpayers adjusting their income around those 'notches.' In fact, foreigners in high tax municipalities tend to lower their income just below the threshold to enjoy the special fiscal regime. Those in low tax municipalities lift their income just above the threshold to get into the ordinary taxation scheme. Given that the special regime expires after five years, the expectation is that high-income foreigners in high-tax municipalities will choose to relocate to low-tax municipalities when the special regime expires, while those located in low-tax municipalities are likely to stay. Using detailed data on location decisions, income, labour market conditions, and socio-demographics from individuals that are eligible for the special tax scheme, they find evidence that local tax differences can induce internal mobility of foreigners that would otherwise experience a tax rise, with the probability of moving being increasing in the expected returns from relocation.

Agrawal and Foremny (2019) study a fiscal decentralization reform in Spain that allowed regions to set their own residence-based personal income tax rates and brackets, and estimate its effect on the mobility of individual taxpayers

during the period 2005-2014. The reform led to large marginal tax rate differentials, particularly on high-income taxpayers, and raised worries that tax-induced mobility of wealthy taxpayers would make redistribution unfeasible. Based on sample data of taxpayers from Social Security records, and using individual movers' data, they find a significant impact of regional taxes on location choices, with a 1 percent increase in the net of tax rate in a region increasing the probability of moving to that region by 1.7 percentage points. In an aggregate region-by-region pairwise analysis that allows estimation of the stock elasticity of the population of top-income taxpayers in a region with respect to its own net-of-average-tax-rate, the elasticity is estimated to be around 0.85. In order to simulate the revenue impact of an income tax policy change that raises the marginal income tax rate on incomes above the highest income bracket while leaving all other rates unchanged, they decompose the total effect into three effects. The first is a mechanical revenue increase from raising the tax rate, the second is a behavioral one coming from changes in taxable income that depends on the taxable income elasticity, and the third is the mobility effect depending on the stock elasticity of migration. They show that the (positive) mechanical effect from raising the top marginal tax rate largely dominates the other two (negative) effects in all Spanish regions, thus making the prospect of a race-to-the-bottom in top income taxation and of unsustainability of local systems of income redistribution implausible.

Finally, Eugster and Parchet (2019) analyze theoretically and empirically the constraints created by tax base mobility between proximate jurisdictions that prevent decentralized governments from selecting their 'culturally desired' fiscal policies. Even if nearby jurisdictions have persistent and measurable differences in preferences over publicly provided goods that should lead to significantly different tax policies, the fact of being spatially close might reduce their tax differentials through competition for mobile and heterogeneous individuals. Empirically, they employ a difference-in-differences approach to identify the existence of tax competition that is based on the comparison of tax differentials between jurisdictions that share a border at which preferences change discontinuously (French-speaking and German-speaking municipalities in bi-lingual cantons in Switzerland) with jurisdictions with analogous differences in preferences, but that are located far apart on opposite sides of the language border. Exploiting the fact that direct democracy institutions are used frequently in Switzerland and allow to elicit people's preferences on a number of crucial issues, they proxy

preferences by the share of support for leftist/progressive referenda. Their results suggest that competition tends to offset culture-related tax differentials: at the language border, the difference-in-difference estimate of the impact of vicinity to the language border on the tax differential is large and highly significant, making the observed tax policies on either side of the border remarkably similar, while the corresponding estimate for the preference proxy differential is zero. Moreover, the difference-in-difference estimate of the border on the tax differential is virtually unaffected when controlling for the preference (vote share) proxies. This suggests that the smooth tax gradient that is observed around the language border cannot be explained by mere people's sorting, but is rather the outcome of strategic behavior of local governments anticipating the sorting of heterogeneous individuals.

3 Multi-level income taxation and estimation of a spatial reaction function

3.1 Theoretical background

Consider a two-tiered structure of government, where the upper-tier (regional) government is indexed by R and produces the public good g_R and the lower-tier (municipal) government is indexed by c and produces the public good g_c. Assume that the preferences of a continuum of households living in locality c can be represented by the strictly quasi-concave utility function $u(i, g_R, g_c)$, where i denotes private consumption. Let the regional and municipal public goods be funded by the income taxes set by each tier on residents' gross income y, with tax rate schedules $\tau_R(y)$ and $\tau_c(y)$ respectively. Let $\overline{\tau}_R(y)$ and $\overline{\tau}_c(y)$ denote the average regional and municipal tax rates on income y. Both schedules might be progressive: $\tau'_R(y), \tau'_c(y) \geq 0$.

Assuming preferences are not homothetic, we follow Basten et al. (2017) and write the indirect utility function of a resident in locality c in region R as:

$$u(i^*, g_R^*, g_c^*) = v(\tau_R, \tau_c, \alpha, \beta, \omega) = \frac{1}{\alpha} y^\alpha \left[1 - \overline{\tau}_R(y) - \overline{\tau}_c(y)\right]^\alpha + \frac{1}{\beta} g_R^\beta + \frac{1}{\omega} g_c^\omega \quad (1)$$

where (i^*, g_R^*, g_c^*) is the ideal consumption bundle, and $\alpha, \beta, \omega \in [0, 1]$ are parameters. It is easy to show that the marginal rate of substitution between each of the public goods and the respective average tax rates $(MRS_{R\overline{\tau}}, MRS_{c\overline{\tau}})$ is positive, suggesting that an increase in taxes must be compensated by a higher

amount of the public good:

$$MRS_{R\overline{\tau}} = \frac{dg_R}{d\overline{\tau}_R} = -\frac{\partial v/\partial \overline{\tau}_R}{\partial v/\partial g_R} = \frac{y^\alpha \left[1 - \overline{\tau}_R(y) - \overline{\tau}_c(y)\right]^{\alpha-1}}{g_R^{\beta-1}} > 0 \qquad (2)$$

$$MRS_{c\overline{\tau}} = \frac{dg_c}{d\overline{\tau}_c} = -\frac{\partial v/\partial \overline{\tau}_c}{\partial v/\partial g_c} = \frac{y^\alpha \left[1 - \overline{\tau}_R(y) - \overline{\tau}_c(y)\right]^{\alpha-1}}{g_c^{\omega-1}} > 0 \qquad (3)$$

Moreover, $MRS_{R\overline{\tau}}$ and $MRS_{c\overline{\tau}}$ are increasing in gross household income, implying that higher income households have a higher willingness to pay in terms of lower consumption of the public good in order to benefit from lower income tax rates:

$$\frac{\partial MRS_{R\overline{\tau}}}{\partial y} = \frac{\alpha y^{\alpha-1} \left[1 - \overline{\tau}_R(y) - \overline{\tau}_c(y)\right]^{\alpha-1}}{g_R^{\beta-1}} \left[1 + \frac{(1-\alpha)y\tau_R'(y)}{\alpha(1 - \overline{\tau}_R(y) - \overline{\tau}_c(y))}\right] > 0$$
$$\qquad (4)$$

$$\frac{\partial MRS_{c\overline{\tau}}}{\partial y} = \frac{\alpha y^{\alpha-1} \left[1 - \overline{\tau}_R(y) - \overline{\tau}_c(y)\right]^{\alpha-1}}{g_R^{\beta-1}} \left[1 + \frac{(1-\alpha)y\tau_c'(y)}{\alpha(1 - \overline{\tau}_R(y) - \overline{\tau}_c(y))}\right] > 0$$
$$\qquad (5)$$

As a result, revenue maximising local authorities might engage in competition for attracting high income taxpayers by lowering the marginal tax rates at the upper end of the income tax schedule. As shown by Agrawal and Foremny (2019), a small perturbation in the top marginal income tax rate will have three effects on revenues: a mechanical effect directly resulting from the tax rate change; a taxable income response from taxpayers that do not move; and a migration effect. As a result, the total revenue effect will be positive or negative depending on the elasticity of taxable income and on the stock elasticity of migration, and their estimates on Spanish autonomous communities suggest that the mechanical effect tends to dominate the behavioral effects.

3.2   Estimation strategy

First, we test on $(N \times T)$ panel data whether the share of high income taxpayers at the municipal level in a two-tiered structure of government with overlapping tax bases (regions-municipalities) is affected by the consolidated net-of-tax rate $1 - \tau$, with $\tau = \tau_c^{\max} + \tau_R^{\max}$ equalling the sum of the top marginal tax rates set by municipal and regional authorities, as in equation (6a):

$$\mathbf{b} = \lambda(1 - \boldsymbol{\tau}) + \mathbf{X}\boldsymbol{\gamma} + \mathbf{u} \qquad (6a)$$

where $\mathbf{b}$ denotes the $(NT \times 1)$ vector of the top income shares measured at the municipal level of government, $\mathbf{X}$ is a matrix of controls, and $\mathbf{u}$ is a $(NT \times$

1) vector residuals that might be spatially correlated. Given that municipal authorities might set their policy based on the characteristics of the taxpayers that are located there, we instrument $(1 - \boldsymbol{\tau})$ by its corresponding top regional retention rate $(1 - \boldsymbol{\tau}_R^{\max})$.

Next, we test for *horizontal fiscal externalities*, that is, we investigate whether the income tax rates of lower-tier neighboring authorities (municipalities sharing a common border) are set interdependently. To do so, we specify the panel data spatial tax reaction function (7), where the $(NT \times 1)$ vector of consolidated tax rates observed at the municipal level of government $(\boldsymbol{\tau})$ is modelled as a first-order spatial auto-regressive process:

$$\boldsymbol{\tau} = \rho \mathbf{W} \boldsymbol{\tau} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (7)$$

where ρ is a scalar coefficient, with $0 < |\rho| < 1$, $\mathbf{W} = [\mathbf{I}_T \otimes \mathbf{W}_N]$ is a block-diagonal, row-standardized matrix of non-stochastic weights built on localities' geographic coordinates, $\mathbf{I}_T$ is the identity matrix of dimension T, $\mathbf{W}_N = \{w_{ij} \geq 0\}$, $i, j = 1, ..., N$, and $\boldsymbol{\varepsilon}$ is a $(NT \times 1)$ vector of residuals[3] In defining the weights w_{ij}, we follow the conventional binary contiguity criterion. Upon matrix-row-normalization, the link between localities i and j (w_{ij}) equals $\frac{1}{n_i}$ if jurisdiction j is adjacent (shares a common border) to jurisdiction i, 0 otherwise, with n_i being the number of units being adjacent to (sharing borders with) unit i.

Equation (7) can be estimated by maximum likelihood methods if one is willing to make distributional assumptions on $\mathbf{u}$, or on moment conditions arising from the hypothesis of strict exogeneity of $\mathbf{X}$ and its spatial transformations $\mathbf{WX}, \mathbf{W}^2\mathbf{X}, ...$ A particularly popular and intuitive application of this principle has been to use the spatially weighted average of first-order neighbors' exogenous variables $\mathbf{WX}$ as instruments for the spatially weighted average of their tax policies $\mathbf{W}\boldsymbol{\tau}$ in a spatial instrumental variable (SIV) approach:

$$\widehat{\boldsymbol{\gamma}}_{SIV} = \widehat{\left[\rho, \boldsymbol{\beta}'\right]}' = \frac{(\mathbf{WX}, \mathbf{X})' \boldsymbol{\tau}}{(\mathbf{WX}, \mathbf{X})' (\mathbf{W}\boldsymbol{\tau}, \mathbf{X})} \qquad (8)$$

Alternatively, if one is not willing to assume that first-order (or higher order) neighbors' characteristics can be legitimately excluded from equation (7) or to make formal distributional assumptions on $\mathbf{u}$, then consistent estimation

---

[3] All variable are taken in deviations from group means, and the matrix $\mathbf{X}$ contains year dummies or region-year dummies.

of $\boldsymbol{\gamma}$ requires an ad hoc instrument matrix $\mathbf{Z}$ such that $E\left(\mathbf{Z}'\mathbf{W}\boldsymbol{\tau}\right) \neq 0$ and $E\left(\mathbf{Z}'\mathbf{u}\right) = 0$.

In this regard, Parchet (2019) considers the concurrent taxation of household income by municipalities and cantons in Switzerland, and proposes an estimator of $\boldsymbol{\gamma}$ that uses information on all municipalities that have at least one 'neighbor' in a different canton (thus removing from the dataset all 'internal' municipalities), and that employs the tax rate set by the 'foreign' canton (call it τ_{Ft}) as an instrument for the spatially weighted average of the neighbors of any given municipality c that is located in region R (equation (9) below):

$$
\begin{aligned}
\tau_{-cRt} &= \sum_{j \in R} w_{cj}\tau_{jRt} + \sum_{k \in F \neq R} w_{ck}\tau_{kFt} \\
&= \sum_{j \in R} w_{cj}\left(\tau_{jt} + \tau_{Rt}\right) + \sum_{k \in F \neq R} w_{ck}\left(\tau_{kt} + \tau_{Ft}\right) \quad (9)
\end{aligned}
$$

where:

$$
\sum_{j \in R} w_{cj} + \sum_{k \in F \neq R} w_{ck} = 1 \quad (10)
$$

$$
w_{cj} = \left\{ \begin{array}{cc} \frac{1}{N_c} & \text{if} \quad j \in R_c \\ 0 & 0 \end{array} \right. ; \quad w_{ck} = \left\{ \begin{array}{cc} \frac{1}{N_c} & \text{if} \quad k \in F_c \\ 0 & 0 \end{array} \right. \quad (11)
$$

$$
N_c = N_{cR} + N_{cF} = \sum_j 1(j \in R_c) + \sum_k 1(k \in F_c) \quad (12)
$$

In particular, in order to allow for the fact that the influence of the foreign canton's rate on the weighted average of neighbors might depend on the relative number of municipalities that are actually located in that foreign canton, Parchet (2019) multiplies the foreign canton's tax rate by the fraction of the neighboring municipalities that are actually located there (equation (13)), leading to instrument z_{cRt} (equation (14)):

$$
s_{cF} = \sum_{k \in F \neq R} w_{ck} = \frac{N_{cF}}{N_c} \quad (13)
$$

$$
z_{cRt} = s_{cF}\tau_{Ft} = \sum_{k \in F \neq R} w_{ck}\tau_{Ft} = \frac{N_{cF}}{N_c}\tau_{Ft} \quad (14)
$$

We first estimate the spatial reaction function (7) by an overidentified spatial instrumental variables approach based on the moment condition $E\left(\mathbf{W}\mathbf{X}'\mathbf{u}\right) = 0$, with the spatial weights matrix defined above. Next, we compare those results to those that can be obtained when employing the just identified border

14

discontinuity IV estimator suggested by Parchet (2019), using either τ_{Ft} or z_{cRt} as instrument for τ_{-cRt}. In fact, though, the spatial reaction function (7) is based on the very restrictive hypothesis that municipal tax rates react in the same way to municipal fiscal policies, to the own region's fiscal policy, and to the fiscal policies of nearby regions. This implies that the validity of the instrument τ_{Ft} (or z_{cRt}) is disputable as well as untestable.

The above approach can be extended by estimating a more flexible specification where the dependent variable is the lower-tier (municipal) tax rate (τ_{ct}), that is allowed to depend on the tax rate of the upper-tier region where municipalities are located (τ_{Rt}), on the weighted average of municipal tax rates in neighboring municipalities, and on the nearby region's tax rate (τ_{Ft}):

$$\tau_{ct} = \underbrace{\delta_R \tau_{Rt}}_{\text{vertical}} + \underbrace{\delta_F \tau_{Ft}}_{\text{diagonal}} + \theta_h \underbrace{\sum_{j \neq c} w_{cj} \tau_{jt}}_{\text{horizontal}} + \mathbf{x}'_{ct} \boldsymbol{\pi} + \zeta_{ct} + \tag{15}$$

Under the testable hypothesis that neighbors' characteristics $\mathbf{x}'_{jt}$ can be excluded from equation (15), then an IV estimator using those variables as instruments for the endogenous spatial lag $\sum w_{cj} \tau_{jt}$ returns consistent estimates of the parameter vector $[\delta_R, \delta_F, \theta_h, \boldsymbol{\pi}']$. Ideal instruments for municipal own tax policies are components of their budget constraints that are transferred to them by upper-level authorities (such as EU, state, or regional lump-sum grants), given that it seems plausible that a lump-sum grant to a locality has no direct influence on other localities than the one that receives it, though it can indeed have an indirect influence insofar as it changes the own fiscal policy and therefore the attractiveness of the recipient authority.

4 Local income taxation in Italy

We perform the empirical analysis on a panel dataset of around 8,000 Italian municipalities that are located in 20 regions. Both upper-level regional authorities and lower-level municipal ones set income tax rates on the same personal income tax base as the national government. Municipal and regional income taxes can either be proportional or progressive. In the former case, the local (municipal, regional) government sets the flat tax rate that applies to taxpayers' gross income and that works as a surcharge on the progressive nationwide schedule. If they opt for progressive taxation, local authorities can set the marginal tax rates applying to the income tax brackets, but the tax brackets must

be identical as the ones set by the national government for its own share of the tax.

Regional income tax rates exceed by large the rates set by municipal governments. The maximum applicable regional rate is 4%, almost five times as large as the municipal one (0.9%). Overall this implies that while within-region differences in income tax burdens end up to be modest, taxpayers can face differences in marginal income tax rates between residences in different regions exceeding four percentage points. Descriptive statistics are reported in table 1.

5   Estimation results

Table 2 reports estimates of the tax base response (share of taxpayers with income above euro 120,000) to the consolidated municipal-regional net-of-tax rate $1 - \boldsymbol{\tau}^{\max} = \mathbf{1} - (\boldsymbol{\tau}_c^{\max} + \boldsymbol{\tau}_R^{\max})$. While the OLS estimates are not significantly different from zero, the IV estimates using the regional net-of-tax rate $(\mathbf{1} - \boldsymbol{\tau}_R^{\max})$ as instrument are positive and statistically significant in column (2). Similar results emerge when using the one-year lag of the regional net-of-tax rate as instrument in column (4).

Table 3 first provides in column (1) OLS estimates of the spatial specification (7) on the entire sample of municipalities for the years 2003-2017, whether or not they are close to a regional border. Column (2) presents OLS estimates on the restricted sample of municipalities having at least one neighbor in another region (1,021 municipalities for about 15,000 observations in this case). Column (3) applies the SIV estimator to that same sample, using first-order neighbors' exogenous characteristics as instruments. In all instances, the variable of interest is the consolidated municipal plus regional maximum income tax rate: $\boldsymbol{\tau}^{\max} = \boldsymbol{\tau}_c^{\max} + \boldsymbol{\tau}_R^{\max}$. This is the top location-dependent marginal tax rate, and varies from 0.9% to 4.5%. The OLS estimates of the key auto-correlation coefficient are around 1 both in the overall sample and in the restricted border one. Similarly, also the SIV approach returns an estimate of around 1.

Table 4 reports the estimates of the same equation when including region-year dummies. This amounts to using the municipal rate as the dependent variable, and returns a much smaller autoregressive coefficient. In this case, the first stage regression using neighbors' characteristics as instruments fails both the tests of instruments' relevance and validity.

Table 5 reports instead the estimates when employing the just identified

estimator using neighboring regions' tax rates as instrument. As in Parchet (2019), the coefficient of neighboring municipalities' tax rates is negative and significant in columns (1) and (3), when those rates are weighted or are not weighted by the share of external neighbors s_{cF}. On the other hand, when the instrument is taken with a one-year lag, the coefficient is not estimated to be different from zero in columns (5) and (7). Finally, the estimated coefficient is not statistically different from zero in all specifications of table 6 that include region-year dummies .

6 Conclusions

This paper has exploited the multi-level structure of progressive personal income taxation in Italy, and particularly the discontinuity of upper-tier top marginal rates at regional borders, to analyze the response of the share of high income taxpayers to local tax differentials and the tax rate setting decisions of lower-tier authorities relative to neighboring ones. We have employed recently proposed border-discontinuity instrumental variable estimators of spatial reaction functions that use cross-border upper-tier tax policies as instruments for lower-tier ones, and have compared them to conventional instrumental variable estimators of first-order spatial auto-correlation parameters. Using a large panel dataset on the distribution of taxable income and on municipal and regional income tax rates spanning through two decades, we have found a positive and significant impact of the net-of-tax rate on the share of top income taxpayers that reside in a locality. Next, when estimating a fiscal reaction function by means of border-discontinuity instrumental variable estimators that use cross-border upper-tier tax policies as instruments for lower-tier ones, we found that the evidence of negative spatial correlation is highly sensitive to the specification of the empirical model, leaving open the issue of whether local income tax rates are strategic complements or substitutes.

References

[1] Agrawal, D., Foremny, D., Relocation of the rich: Migration in response to top tax rate changes from Spanish reforms, Review of Economics and Statistics (2019) forthcoming.

[2] Basten, C., Ehrlich, M., Lassmann, A., Income taxes, sorting, and the costs of housing: Evidence from municipal boundaries in Switzerland, Economic Journal 127 (2017) 653-687.

[3] Di Porto E., Revelli, F., Tax limited reaction functions, Journal of Applied Econometrics 28 (2013) 823-839.

[4] Eugster, B., Parchet, R., Culture and taxes, Journal of Political Economy (2019) forthcoming.

[5] Gibbons S., Overman H., Mostly pointless spatial econometrics? Journal of Regional Science 52 (2012) 172-191.

[6] Isen, A., Do local government fiscal spillovers exist? Evidence from counties, municipalities, and school districts, Journal of Public Economics 110 (2014) 57-73.

[7] Liebig, T., Sousa-Poza, A., The influence of taxes on migration: Evidence from Switzerland, Cambridge Journal of Economics 30 (2006) 235-252.

[8] Liebig, T., Puhani, P., Sousa-Poza, A., Taxation and Internal Migration: Evidence from the Swiss Census Using Community-Level Variation in Income Tax Rates, Journal of Regional Science 47 (2007) 807-836.

[9] Manski, C., Identification of endogenous social effects: The reflection problem, Review of Economic Studies 60 (1993) 531-542.

[10] Martinez, I., Beggar-thy-neighbour tax cuts: Mobility after a local income and wealth tax reform in Switzerland, Working Paper LISER No. 08 (2017).

[11] Parchet, R., Are local tax rates strategic complements or strategic substitutes? American Economic Journal: Economic Policy (2019) forthcoming.

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

[13] Schaltegger, C., Somogyi, F., Sturm, J.-E., Tax competition and income sorting: Evidence from the Zurich metropolitan area, European Journal of Political Economy 27 (2011) 455-470.

[14] Schmidheiny, K., Income segregation and local progressive taxation: Empirical evidence from Switzerland, Journal of Public Economics 90 (2006) 429-58.

[15] Schmidheiny, K., Slotwinski, M. Behavioral responses to local tax rates: Quasi-experimental evidence from a foreigners' tax scheme in Switzerland, Journal of Public Economics 167 (2018) 293-324.

[16] Young, C., Varner, C., Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment, National Tax Journal 64 (2011) 255-283.

[17] Young, C., Varner, C., Lurier, I., Prisinzano, R.. Millionaire Migration and Taxation of the Elite: Evidence from Administrative Data, American Sociological Review 81 (2016) 421-446.

Tables

Table 1 Descriptive statistics

	obs.	mean	std. dev.	min	max
Maximum municipal tax rate	118,785	0.385	0.280	0	0.9
Maximum regional tax rate	118,785	1.543	0.537	0.9	4
Share with personal income higher than 120,000 €	110,866	0.003	0.004	0	0.098
Share of over 65 years old males	118,785	0.098	0.031	0	0.75
Share of over 65 years old females	118,785	0.129	0.037	0.01	0.71
Share of under 5 years old males	118,785	0.021	0.006	0	0.22
Share of under 5 years old females	118,785	0.020	0.006	0	0.45
Share of males between 5 and 14 years old	118,785	0.045	0.011	0	0.25
Share of females between 5 and 14 years old	118,785	0.043	0.010	0	0.30

Table 2 TAX BASE AND LOCAL TAX POLICY

	OLS	IV(1)	First stage(1)	IV(2)	First stage(2)
	(1)	(2)	(3)	(4)	(5)
			% taxpayers > 120,000		
$1 - \tau_{cRt}$	0.003	0.007***		0.009***	
	(0.002)	(0.002)		(0.003)	
$1 - \tau_{Rt}$			0.998***		
			(0.002)		
$1 - \tau_{cRt-1}$					0.783***
					(0.006)
Municipality fixed effects			Yes		
Year fixed effects			Yes		
Number of observations			110,866		102,947
Number of municipalities			7,919		7,919
Period			2003-2016		2004-2016

Note: ***p<0.01, **p<0.05, and *p<0.10. The dependent variable is the share of high income (>120,000 euro) taxpayers. All estimations include the share of over 65 years old male and female individuals, the share of under 5 years old male and female individuals, and the share of male and female individuals between 5 and 14 years old. Standard errors are clustered at municipal level.

Table 3 SPATIAL CORRELATION IN TAX RATES

	All municipalities	Border municipalities	
	OLS	OLS	IV
	(1)	(2)	(3)
τ_{-cRt}	0.987***	1.072***	0.977***
	(0.002)	(0.022)	(0.079)
First-stage F-test on instruments			15.13
Over-identification test (p-value)			0.362
Municipality fixed effects		Yes	
Year fixed effects		Yes	
Number of observations	118,785	15,315	15,315
Number of municipalities	7,919	1,021	1,021
Period	2003-2017	2003-2017	2003-2017

Note: ***p<0.01, **p<0.05, and *p<0.10. The dependent variable is the maximum consolidated regional and municipal tax rate on personal income. Border municipalities have at least one neighboring municipality located in another region. All estimations include the share of over 65 years old male and female individuals, the share of under 5 years old male and female individuals, and the share of male and female individuals between 5 and 14 years old. The instruments in column (3) are the average value of controls in neighboring municipalities. The over-identification test reports the p value of the Hansen J statistic. Standard errors are clustered at municipal level.

Table 4 SPATIAL CORRELATION IN TAX RATES

	All municipalities	Border municipalities	
	OLS	OLS	IV
	(1)	(2)	(3)
τ_{-cRt}	0.179***	0.057***	0.640***
	(0.011)	(0.016)	(0.240)
First-stage F-test on instruments			2.66
Over-identification test (p-value)			0.010
Municipality fixed effects		Yes	
Region-year fixed effects		Yes	
Number of observations	118,785	15,315	15,315
Number of municipalities	7,919	1,021	1,021
Period	2003-2017	2003-2017	2003-2017

Note: ***p<0.01, **p<0.05, and *p<0.10. The dependent variable is the maximum consolidated regional and municipal tax rate on personal income. Border municipalities have at least one neighboring municipality located in another region. All estimations include the share of over 65 years old male and female individuals, the share of under 5 years old male and female individuals, and the share of male and female individuals between 5 and 14 years old. The instruments in column (3) are the average value of controls in neighboring municipalities. The over-identification test reports the p value of the Hansen J statistic. Standard errors are clustered at municipal level.

Table 5 STRATEGIC INTERACTION IN TAX RATES: IV STRATEGY

	Border municipalities			
	IV(1)	First stage(1)	IV(2)	First stage(2)
	(1)	(2)	(3)	(4)
τ_{-cRt}	-0.147**		-0.120**	
	(0.064)		(0.059)	
τ_{Ft}		0.313***		
		(0.015)		
z_{cRt}				0.789***
				(0.029)
	IV(3)	First stage(3)	IV(4)	First stage(4)
	(5)	(6)	(7)	(8)
τ_{-cRt}	-0.111		-0.094	
	(0.074)		(0.070)	
τ_{Ft-1}		0.256***		
		(0.014)		
z_{cRt-1}				0.639***
				(0.029)
Municipality fixed effects			Yes	
Year fixed effects			Yes	
Number of observations	15,315	15,315	15,315	15,315
Number of municipalities	1,021	1,021	1,021	1,021
Period	2003-2017	2003-2017	2003-2017	2003-2017

Note: ***p<0.01, **p<0.05, and *p<0.10. The dependent variable is the maximum consolidated regional and municipal tax rate on personal income. Border municipalities have at least one neighboring municipality located in another region. All estimations include the share of over 65 years old male and female individuals, the share of under 5 years old male and female individuals, and the share of male and female individuals between 5 and 14 years old. The instruments in column (3) are the average value of controls in neighboring municipalities. The over-identification test reports the p value of the Hansen J statistic. Standard errors are clustered at municipal level.

Table 6 STRATEGIC INTERACTION IN TAX RATES: IV STRATEGY

	Border municipalities			
	IV(1)	First stage(1)	IV(2)	First stage(2)
	(1)	(2)	(3)	(4)
τ_{-cRt}	-0.003		-0.018	
	(0.022)		(0.023)	
τ_{Ft}		0.348***		
		(0.012)		
z_{cRt}				0.797***
				(0.023)
	IV(3)	First stage(3)	IV(4)	First stage(4)
	(5)	(6)	(7)	(8)
τ_{-cRt}	-0.010		-0.023	
	(0.028)		(0.030)	
τ_{Ft-1}		0.265***		
		(0.009)		
z_{cRt-1}				0.616***
				(0.022)
Municipality fixed effects			Yes	
Year-region fixed effects			Yes	
Number of observations	15,315	15,315	15,315	15,315
Number of municipalities	1,021	1,021	1,021	1,021
Period	2003-2017	2003-2017	2003-2017	2003-2017

Note: ***p<0.01, **p<0.05, and *p<0.10. The dependent variable is the maximum consolidated regional and municipal tax rate on personal income. Border municipalities have at least one neighboring municipality located in another region. All estimations include the share of over 65 years old male and female individuals, the share of under 5 years old male and female individuals, and the share of male and female individuals between 5 and 14 years old. The instruments in column (3) are the average value of controls in neighboring municipalities. The over-identification test reports the p value of the Hansen J statistic. Standard errors are clustered at municipal level.