Strategic interaction in Swiss local income tax setting revisited

Fabian Muny and Lukas Peham

Master Course in Spatial Economics
Winter Term 2019/20 - WU Vienna

February 20, 2020

Abstract

This report assesses the nature of strategic interaction in income tax setting between Swiss municipalities. Using a variety of spatial estimation strategies we find evidence that tax rates are strategic substitutes. Hence, we can reject the hypothesis that harmful tax competition between municipalities led to the decline of tax rates observed during the analysed time period. In addition, we show that for the application at hand the spatial panel maximum likelihood (SPML) estimation reveals insights that point into the same direction as the causal IV method proposed by Parchet (2019a).

Keywords: Income Taxation, Tax Competition, Spatial Econometrics, Regional Economics

1. Introduction

Tax competition is an intensively discussed issue among both economists and policy makers. The direction and degree of strategic interaction in tax setting within and across different levels of government is hard to quantify. Parchet (2019a) contributes to this debate by taking a deep look into the income tax setting behaviour of Swiss municipalities. Switzerland is a very federalist state where local jurisdictions enjoy large fiscal autonomy. For instance, municipalities have the power to set their own personal income tax rates. While this power slightly varies from canton to canton, a Swiss resident in general faces each year the following consolidated income tax rate

$$T_{ic,t} = t_{ic,t} + t_{c,t} \tag{1}$$

consisting of a part determined by the municipality $t_{ic,t}$ and a part specified by the canton $t_{c,t}$. Due to this legal structure, Swiss residents have incentives to change their residence in order to reduce their tax burden. Through this channel it seems plausible that the municipalities' decisions about their tax rate influence each other. This observation brings up the research questions whether and how municipalities interact in personal income tax rate setting.

In game-theoretic terms strategic interaction is classified in two categories. On the one hand, decisions are called strategic complements if they mutually reinforce one another. In the context of local tax competition this would imply that a municipality lowers its tax rate if it expects that competing municipalities do the same. This would support the "race to the bottom" hypothesis according to which the mobility of firms and people leads to harmful tax competition. On the other hand, decisions are termed as strategic substitutes if they mutually offset one another. This would imply that municipalities increase their tax rate when others decrease their tax rates. This type of strategic interaction could be explained by long-term expenditures for public goods. For example, a decrease in the tax base due to lower taxes in other jurisdictions might force a municipality to increase its tax rate in order to be able to cover its cost.

The main result of Parchet (2019a) is that municipality's personal income tax rates are indeed strategic substitutes. He arrives at this conclusion by applying an estimation approach which uses neighboring cantons' tax rates for border municipalities as instruments. Based on the critique by Gibbons and Overman (2012), he argues that, in contrast to alternative spatial econometric methods, this approach enables causal inference.

In this paper we replicate the baseline results of Parchet (2019a) and contribute to the scientific debate by critically discussing the identification methods available, by examining the robustness of the findings, and by applying alternative spatial estimation methods. In particular, we extend the spatial two-stage least squares (S2SLS) approach by adding more spatial lags to the instrument and apply spatial panel maximum likelihood (SPML) estimation. Our findings provide evidence that local Swiss personal income tax rates are strategic substitutes, which supports the IV findings of Parchet (2019a).

The outline of the paper is as follows. Section 2 discusses the data and provides a brief descriptive analysis. Section 3 explains the methods used in the replication and the extension. Finally, section 4 presents the results before we conclude the analysis in section 5.

2. Data and Descriptive Analysis

The analysis in this paper is based on the original data set provided by Parchet (2019b). It is an unbalanced panel of 2,428 Swiss municipalities from 1983 until 2012. The main variable of interest is the the consolidated personal income tax rate of an unmarried taxpayer with a gross annual income of CHF 1,000,000. As control variables it covers population, the percentage of foreign nationals, the percentage of young aged less than 20 and old aged more than 80, the percentage of employees in the secondary and tertiary sectors, total full-time equivalent employment per capita, the unemployment rate, the share of votes in favor of left-of-center parties in national elections, and the number of movie theaters within 10 km.

The main explanatory variable of the analysis is the spatial lag of a municipality's consolidated personal income tax rate $\boldsymbol{W}T_{ic,t}$. In order to obtain this variable we need to define a weights matrix $\boldsymbol{W}$. We follow the baseline approach of Parchet (2019a) and define the weights by a 10km road distance cutoff:

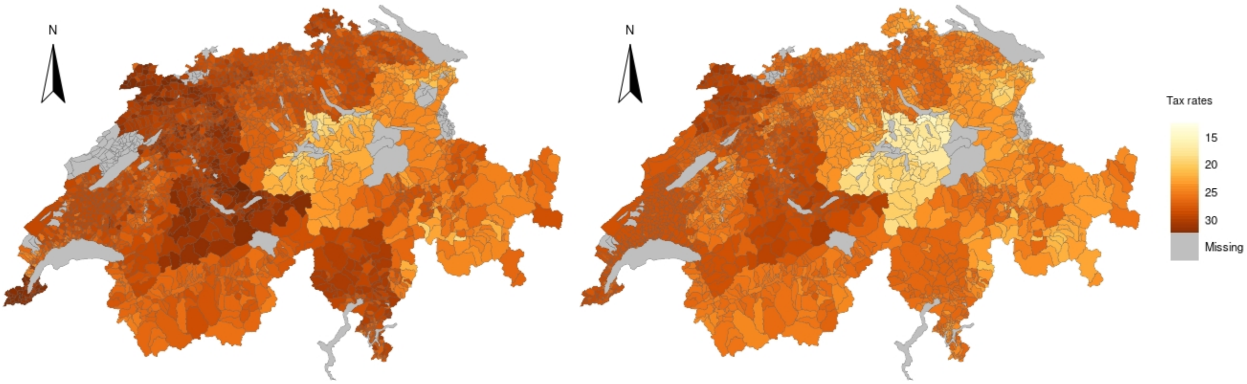
$$w_{ij} = \begin{cases} 1/N, & \text{if } d_{ij} \leq 10 \\ 0, & \text{otherwise} \end{cases}, \quad N = \sum_{j \neq i} \mathbf{1}_{d_{ij} \leq 10}$$

Following this equation, municipality i is a neighbor of municipality j if the shortest road distance d_{ij} between the two municipalities is less or equal to 10km. Since we assign uniform weights to each neighbor, $\boldsymbol{W}T_{ic,t}$ can be interpreted as the average tax rate of a municipality's neighbors.[1]

In the course of the analysis we are using two different samples of the full data set. The first one we call the *border sample*. This sample consists of all municipalities that have at least one neighbor in a different canton as defined by the weights matrix $\boldsymbol{W}$. It covers 1,047 municipalities for the whole observation period. The second sample is the largest possible *balanced subsample* of the provided data set. It is required since spatial panel methods in R (Millo et al., 2012) are only implemented for balanced panels. To balance the data set we first delete 99 municipalities which have data for only twelve years. Next, we exclude the years 1983 to 1985 since there exists no data for 289 municipalities in this time period. Finally,

[1]We sincerely thank Raphaël Parchet for providing us with the confidential road distance data.

Figure 1: Consolidated personal income tax rates of Swiss municipalities in 1990 (left) and 2010 (right). Grey areas are municipalities with missing data or lakes.



we eliminate seven municipalities which have no data for 2011 and 2012. Consequently, we end up with a balanced panel from 1986 to 2012 including 2283 municipalities. Summary statistics of all samples used can be found in table A.3.

Before turning to the estimation methods and results we want to provide a first look into the data to further motivate the relevance of our research question. In figure 1 we display the consolidated personal income tax rates across Swiss municipalities for the years 1990 and 2010. A comparison of the two figures reveals that over the time period of 20 years the consolidated tax rates declined in the majority of the municipalities.[2] Following this observation it is of central interest to find out what factors caused this reduction. As discussed above, one reason could be that municipal tax rates are strategic complements. This would imply that tax competition among municipalities led to the observed decline in tax rates. As we show in the following sections, this explanation is unlikely to hold.

[2]This finding does not only hold for the representative unmarried taxpayer with a gross annual income of CHF 1,000,000 but also for other categories of tax payers and income classes

3. Methods

The empirical analysis of this paper is based on the regression function

$$T_{ic,t} = \rho \boldsymbol{W} T_{ic,t} + \boldsymbol{X}_{ic,t}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_{ic,t}\boldsymbol{\gamma} + \alpha_i + \delta_{c,t} + \epsilon_{ic,t} \tag{2}$$

where $T_{ic,t}$ denotes the consolidated tax rate of municipality i in canton c at time t defined in equation (1). $\boldsymbol{W}T_{ic,t}$ represents the spatial lag of this tax rate which corresponds to the weighted average of neighboring municipalities' consolidated tax rates. Due to the time dimension of our data, the weights matrix presented in section 2 appears as a block matrix $\boldsymbol{W}_{\otimes} = \boldsymbol{I} \otimes \boldsymbol{W}$. In addition, we include a vector of control variables $\boldsymbol{X}_{ic,t}$, the spatial lag of these control variables $\boldsymbol{W}\boldsymbol{X}_{ic,t}$ as well as municipality and canton-year fixed effects α_i and $\delta_{c,t}$. This specification is equivalent to the one in Parchet (2019a).

Due to endogeneity problems caused by the lagged dependent variable, estimates and standard errors of model (2) are inconsistent when estimated by OLS. The main issue is that, by construction, the spatial lag $\boldsymbol{W}T_{ic,t}$ is linked to the error term $\epsilon_{ic,t}$ which violates the exogeneity assumption. For instance, this leads to a problem of reverse causation since $\boldsymbol{W}T_{ic,t}$ itself depends on $T_{ic,t}$. In the following we describe the three alternative estimation methods used in this paper to overcome these obstacles.

3.1. Spatial two-stage least squares estimation (S2SLS)

The first approach has been proposed by Kelejian and Prucha (1998) and applies a two-stage least squares estimation to deal with endogeneity. In particular, it uses spatial lags of the exogenous explanatory variables as instruments for the lagged dependent variable. Translated to model (2) this requires the following procedure: In the first stage, we regress the spatial lag of the consolidated income tax rate $\boldsymbol{W}T_{ic,t}$ on one or more spatial lags of the control variables $\boldsymbol{W}\boldsymbol{X}_{ic,t}, \boldsymbol{W}^2\boldsymbol{X}_{ic,t}, [...]$. Subsequently, we estimate model (2) by OLS replacing $\boldsymbol{W}T_{ic,t}$ with the fitted values of the first-stage regression and omitting $\boldsymbol{W}\boldsymbol{X}_{ic,t}$.

The decision to choose the spatial lags of $\boldsymbol{X}_{ic,t}$ as instruments is motivated by the reduced form of equation (2) which reads as

$$T_{ic,t} = \boldsymbol{X}_{ic,t}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_{ic,t}(\rho\boldsymbol{\beta} + \boldsymbol{\gamma}) + \boldsymbol{W}^2\boldsymbol{X}_{ic,t}(\rho^2\boldsymbol{\beta} + \rho\boldsymbol{\gamma}) + [...] + \tilde{\alpha}_i + \tilde{\delta}_{c,t} + \tilde{\epsilon}_{ic,t} \tag{3}$$

using a Neumann series. Two conditions need to be fulfilled for instrument validity. Firstly, neighboring municipalities' controls should be correlated to the respective consolidated tax rates. This holds by construction. Secondly, the spatial lags of $\boldsymbol{X}_{ic,t}$ may affect municipality i's tax rate only through $\boldsymbol{W}T_{ic,t}$. This assumption is hard to justify since we originally assume that $\boldsymbol{W}\boldsymbol{X}_{ic,t}$ is included in model 2.

As argued in Gibbons and Overman (2012), even if the the exclusion of $\boldsymbol{W}\boldsymbol{X}_{ic,t}$ can be credibly motivated, identification critically depends on a "correct" choice of the weights matrix $\boldsymbol{W}$. For instance, assume tax rates of neighbors within 20km (instead of 10km) road distance affect $T_{ic,t}$. Then, causal identification of ρ breaks down since higher order lags might affect $T_{ic,t}$ directly. Finally, another key problem of S2SLS is that the spatial lags $\boldsymbol{W}\boldsymbol{X}_{ic,t}$, $\boldsymbol{W}^2\boldsymbol{X}_{ic,t}$, [...] tend to be highly correlated which is referred to as a *weak identification problem* that leads to imprecise estimates in the second stage of the procedure.

3.2. Instrumental variable estimation proposed in Parchet (2019a) (IVP)

Motivated by the critique of the S2SLS technique, Parchet (2019a) developed an alternative instrumental variables method that can be applied in the context of Swiss income taxation. He proposes to use neighboring cantons' tax rates of the previous year, $t_{c,t-1}$, as instruments which leads to a first stage regression of the form

$$\boldsymbol{W}T_{ic,t} = \eta(s_{-c}\boldsymbol{W}_{-c}t_{c,t-1}) + u_{ic,t} \tag{4}$$

where $\boldsymbol{W}_{-c}$ denotes a weights matrix assigning a weight of one to municipalities in a different canton within 10km road distance and s_{-c} indicates the share of neighboring municipalities located in another canton.

According to Parchet (2019a), the main advantage of his approach is that it enables to isolate exogenous variations in the tax rates of neighboring municipalities allowing to identify their causal effect. However, the exogeneity argument requires that tax rates of neighboring states, $t_{-c,t-1}$, influence $T_{ic,t}$ only through the spatial lag of $T_{ic,t}$. This implies, inter alia, that taxpayers react in the same way to changes in the canton-level vs. municipality-level tax rates. In addition it assumes that single municipalities do not systematically influence cantonal tax policies and that tax reforms in the neighboring cantons are not driven by unobserved factors which affect $T_{ic,t}$. Moreover, the relevance condition requires that the tax rates of neighboring cantons are correlated with the spatial lag of $T_{ic,t}$. This holds whenever municipalities do not exactly offset the change in $t_{-c,t-1}$.

A disadvantage of the method is that municipalities without a neighbor in a different canton drop out of the sample and cannot be used when applying the method. To justify the representativeness of the remaining observations, Parchet (2019a) shows that border and non-border municipalities share similar population characteristics. Another shortcoming of the method is that it depends on the features of the special Swiss tax system which makes it hard to apply to other settings in the realm of income taxation.

3.3. Spatial panel maximum likelihood estimation (SPML)

The third estimation strategy, SPML, has not been applied in Parchet (2019a). In contrast to OLS, it can be shown that maximum likelihood estimation of spatial models with lagged dependent variables is unbiased and consistent. Following the usual approach in spatial econometrics, we test our model for spatial dependence in the error and the dependent variable using lagrange multiplier tests. As suggested by the test results, we add a spatial lag of the error, $\epsilon_{ic,t} = \boldsymbol{W}\epsilon_{ic,t}\lambda + \nu_{ic,t}$ to model (2). In addition, the tests confirm the use of $\boldsymbol{W}T_{ic,t}$ as explanatory variable. Finally, we perform a Hausman specification test for spatial panel data models that validates the use of fixed effects estimation.

While spatial maximum likelihood estimation is preferred over OLS and widely used by spatial econometricians, it does not come without problems. Gibbons and Overman (2012) argue that the method yields consistent coefficients only under the condition that the estimated model represents the true data generating process. This assumption, however, is questionable in the spatial context since all spatial models are nested within each other. Therefore, without a source of exogenous variation, identification is only possible for the total effect of neighbors' characteristics but one cannot tell whether they are driven by exogenous or endogenous variables.

Despite these shortcomings, we present SPML estimates based on Millo et al. (2012) in the following section and compare them to the results obtained by S2SLS and IVP. As we find, SPML might still yield interesting insights on spatial relationships, especially if *experimentalist* approaches as the one adopted by Parchet (2019a) are not available.

4. Results

4.1. Replication

Table 1 shows several results of model (2) estimated by OLS and S2SLS with the first spatial lag of $\boldsymbol{X}_{ic,t}$ as instrument. The results are presented to gauge the quality of our replication but they do not allow for a causal interpretation. As discussed in the previous section, OLS is inconsistent due to endogeneity and S2SLS is likely to be based on an invalid instrument.

Table 1: Estimation results of model (2) using different samples, weights matrices and estimation techniques. Control variables included but not shown. The full results can be found in table A.4. Standard errors are clustered by municipality and year.

	Dependent variable: $T_{ic,t}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Method	OLS	S2SLS(1)	OLS	S2SLS(1)	OLS	S2SLS(1)	OLS	S2SLS(1)
Sample	full sample		balanced sample		balanced sample		border sample	
$\boldsymbol{W}$	original $\boldsymbol{W}$		original $\boldsymbol{W}$		replicated $\boldsymbol{W}$		original $\boldsymbol{W}$	
$\boldsymbol{W}T_{ic,t}$	0.444***	0.613***	0.457***	0.676***	0.455***	0.684***	0.425***	0.662***
	(0.031)	(0.124)	(0.033)	(0.130)	(0.032)	(0.134)	(0.045)	(0.166)
Observations	70,127	70,127	61,641	61,641	61,641	61,641	28,362	28,362
Adjusted R^2	0.963	0.963	0.964	0.963	0.964	0.963	0.969	0.968
$\boldsymbol{W}\boldsymbol{X}$	NO	NO	NO	NO	NO	NO	NO	NO
Note:						*p<0.05; **p<0.01; ***p<0.001		

The coefficients in columns (1) and (2) are estimated on the full sample from 1983 to 2012 using the weighted spatial lags provided in the original data set (Parchet, 2019b). The OLS estimate and the standard error in column (1) is exactly the same as the coefficient presented in Parchet (2019a, table 2). This comes at no surprise since we are using the same data and the same methods. Thus, we show that the original results computed in Stata can be replicated in R using the lfe package (Gaure, 2013).

Columns (3) and (4) of table 1 present the results for the balanced sample which is not used in Parchet (2019a). A comparison of the coefficients with the first two columns shows that the loss of 12% of the observations does not alter the outcomes and indicates that there are no problems with sample selection. This ensures comparability of our extension based on the balanced sample with the original results in Parchet (2019a).

The estimates in columns (5) and (6) are based on the same sample and the same methods as the previous two columns. However, now the spatial lags are computed with our replicated weights matrix. The change in $\boldsymbol{W}$ leads only to minor changes in the coefficients and standard errors. The reason for the difference is that the weighted spatial lags from the original data set used in columns (3) and (4) have been computed using the full sample of 2,428 municipalities while our spatial lags take only the 2,283 municipalities into account that remain in the balanced sample. Due to the small differences in the coefficients we are optimistic that the loss of municipalities in $\boldsymbol{W}$ does not distort the outcomes.

In a next step, columns (7) and (8) present OLS and S2SLS estimates for the border sample that is subsequently used for the IVP method. Again, the estimates and standard errors correspond exactly to Parchet (2019a, table 2). Even though the border sample consists of less than half of the observations compared to the full sample, the columns show only minor differences. Hence the regression outcomes show great robustness with respect to the underlying sample used in the estimation.

Finally, column (1) of table 2 shows our replication of the IVP approach. Instead of the spatial lags of $\boldsymbol{X}_{ic,t}$ we use neighboring cantons' tax rates of the previous year, $t_{c,t-1}$, as instruments. The coefficient of -0.497 coincides with Parchet (2019a, table 3) and is the key finding of the author. As he argues, his method is able to identify the causal effect of neighboring municipalities' income tax rates on a municipality's own tax rate while the other two methods are not able to do so due to conceptual issues and doubtful assumptions. Hence, he concludes that local income tax rates in Switzerland are strategic substitutes.

Following the discussion of the methods in chapter 3 we share Parchet's doubts on the ability of "conventional spatial methods" to identify causal effects. However, in the next section we show, at least for this application, that the criticised spatial approaches might still reveal insights that point into the right direction.

4.2. Extensions

Our extensions to Parchet (2019a) are threefold: Firstly, we verify the robustness of IVP with respect to the exclusion of the control variables' spatial lags. Secondly, we add spatial lags to the S2SLS instrument and check whether the outcomes change. Finally, we estimate model (2) using SPML. The results are shown in table 2, columns (2) to (7).

Table 2: Estimation results of model (2) using different samples, weights matrices and estimation techniques. Spatial lags of the control variables and the error are not shown. Standard errors are clustered by municipality and year in columns (1) to (5).

	Dependent variable: $T_{ic,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Method	IVP	IVP	S2SLS(1)	S2SLS(2)	S2SLS(3)	SPML	SPML
Sample	border sample		balanced sample			balanced sample	
$\boldsymbol{W}$	original $\boldsymbol{W}$		replicated $\boldsymbol{W}$			replicated $\boldsymbol{W}$	
$\boldsymbol{W}T_{ic,t}$	-0.497^{*}	-0.478^{*}	0.684^{***}	0.680^{***}	0.724^{***}	-0.606^{***}	-0.595^{***}
	(0.223)	(0.232)	(0.134)	(0.123)	(0.108)	(0.009)	(0.009)
Population	−0.128	−0.080	-0.039^{*}	-0.039^{*}	-0.039^{*}	-0.021^{***}	-0.025^{***}
	(0.063)	(0.061)	(0.015)	(0.015)	(0.015)	(0.007)	(0.006)
% foreign	0.025^{*}	0.031^{**}	0.008	0.008	0.008	0.010^{***}	0.009^{***}
	(0.009)	(0.010)	(0.005)	(0.005)	(0.005)	(0.002)	(0.001)
% young	0.006	0.003	0.015^{*}	0.015^{*}	0.015^{*}	0.017^{***}	0.013^{***}
	(0.008)	(0.009)	(0.006)	(0.006)	(0.006)	(0.002)	(0.001)
% old	−0.023	−0.033	0.011	0.011	0.011	0.020^{***}	0.014^{***}
	(0.021)	(0.022)	(0.010)	(0.010)	(0.010)	(0.004)	(0.003)
% secondary	-0.029^{***}	-0.035^{***}	-0.021^{***}	-0.021^{***}	-0.021^{***}	-0.023^{***}	-0.020^{***}
	(0.006)	(0.007)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)
% tertiary	-0.019^{**}	-0.026^{**}	-0.018^{***}	-0.018^{***}	-0.017^{***}	-0.019^{***}	-0.016^{***}
	(0.006)	(0.007)	(0.004)	(0.003)	(0.003)	(0.001)	(0.001)
Employment p.c.	−0.445	−0.342	-0.331^{*}	-0.331^{*}	-0.331^{*}	-0.355^{***}	-0.311^{***}
	(0.472)	(0.495)	(0.131)	(0.131)	(0.130)	(0.045)	(0.040)
% unemployed	0.023	0.024	0.016	0.016	0.016	0.015^{***}	0.015^{***}
	(0.019)	(0.020)	(0.009)	(0.009)	(0.009)	(0.003)	(0.002)
% voting left	−0.005	−0.007	−0.003	−0.003	−0.003	-0.007^{***}	-0.004^{***}
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
# cinemas	−0.012	−0.001	0.012^{*}	0.012^{*}	0.012^{*}	0.018^{***}	0.027^{***}
	(0.015)	(0.011)	(0.006)	(0.006)	(0.006)	(0.003)	(0.003)
Observations	28,362	28,362	61,641	61,641	61,641	61,641	61,641
Adjusted R^2	0.957	0.956	0.963	0.963	0.963		
$\boldsymbol{W}\boldsymbol{X}_{ic,t}$	YES	NO	NO	NO	NO	YES	NO
$\boldsymbol{W}\boldsymbol{\epsilon}_{ic,t}$	NO	NO	NO	NO	NO	YES	YES
Note:							*p<0.05; **p<0.01; ***p<0.001

In Parchet (2019a), the author is inconsistent in the use of the spatial lag of the controls as explanatory variables. While he does not include $\boldsymbol{W}\boldsymbol{X}_{ic,t}$ in the OLS regressions, he does so in his IV approach. Of course, he does not include the spatial lags in the S2SLS estimations since they are serving as instruments and this would violate the exogeneity condition. In order to make the results of the models comparable we estimate the IVP without $\boldsymbol{W}\boldsymbol{X}_{ic,t}$ in column (2) of table 2. As the coefficient shows, the exclusion of the spatial lag does not alter the outcomes.

Columns (3) to (5) of table 2 show estimations using S2SLS with different compositions of the instrument. As a reference, column (3) restates the S2SLS estimate instrumented with the first spatial lag of $\boldsymbol{X}_{ic,t}$ shown in column (6) of table 1. As explained in section 3.1 the choice of $\boldsymbol{W}\boldsymbol{X}_{ic,t}$ is motivated by the reduced form of our model. The reduced form shown in equation (3), however, consists not only of the first lag of $\boldsymbol{X}_{ic,t}$ but of the infinite series of lags of $\boldsymbol{X}_{ic,t}$, namely $\boldsymbol{W}\boldsymbol{X}_{ic,t}, \boldsymbol{W}^2\boldsymbol{X}_{ic,t}, [...]$. Hence, one can argue that instrumenting only with $\boldsymbol{W}\boldsymbol{X}_{ic,t}$ as done in Parchet (2019a) is not enough.

Columns (2) and (3) represent the S2SLS estimates with two and three lags of $\boldsymbol{X}_{ic,t}$, respectively. As the results show, an increase in the number of spatial lags in the instrument does not lead to big changes in the coefficients. If anything, the size of the estimates increases but they do not come closer to the IVP result. Therefore, we can conclude that the S2SLS estimate is robust to the number of lags it is instrumented with and that the distortions in the outcome do not diminish when increasing the number of lags.

Finally, the last two columns show the results of the estimation by SPML. While column (6) includes a lag of $\boldsymbol{X}_{ic,t}$, it is omitted in column (7). Interestingly, the coefficients of $\boldsymbol{W}T_{ic,t}$ are negative and comparable in size to the results estimated by IVP. Again, there are little differences in our coefficient of interest between the specifications including and excluding the spatial lag of $\boldsymbol{X}_{ic,t}$. In addition, the coefficients of the control variables are very similar to the ones shown for the S2SLS regressions which strengthens the cogency of our results. The differences are slightly larger compared to the IVP coefficients. However, this might be due to the smaller sample size used for the IVP regressions.

Note that the standard errors in columns (6) and (7) are not clustered since this is not (yet) possible within the splm package in R (Millo et al., 2012). As a consequence, the significance levels of the estimates are not reliable as there might be correlations in the error along the cross-sectional and time dimensions. However, due to the small standard error in comparison to the coefficient we are optimistic that the estimate would remain significant even after clustering the standard errors.

The finding of a negative coefficient on $\boldsymbol{W}T_{ic,t}$ in the SPML regression allows for two conclusions. On the one hand, it provides evidence that Swiss municipalities' income tax rates are strategic substitutes. This outcome coincides with Parchet (2019a) and shows that the decline in income tax rates observed in section 2 does not result from harmful competition between municipalities. On the other hand, it contradicts the finding of Parchet (2019a) that the outcomes of standard spatial econometric approaches are not in line with his results. In contrast, we show that in the context of this paper SPML yields similar results to the proposed IV method.

5. Conclusion

The preceding analysis sheds light on the nature of strategic interaction in local tax setting in Switzerland. While previous studies using the standard spatial econometric toolkit mainly found that taxes are strategic complements, Parchet (2019a) detects the opposite for Swiss local personal income taxes. On the basis of this surprising finding we contributed to the debate by evaluating the pros and cons of several identification methods, by verifying the robustness of Parchet (2019a) and by applying alternative spatial estimation methods.

The main outcome of this paper is that the results of our SPML estimation support the previous findings of Parchet (2019a) which strengthens the evidence that Swiss municipalities' income tax rates are strategic substitutes. Hence, we can reject the hypothesis that tax competition between municipalities led to the decline of tax rates in the analysed time period. In addition, we have been able to confirm the robustness of the S2SLS and IV approaches applied in Parchet (2019a).

Nevertheless, the findings are not as clear as they appear to be. First of all, the result of no "race to the bottom" in income tax setting between Swiss municipalities is not generalizable to other levels of government. For instance, it might well be the case that income tax competition in Switzerland takes place at the cantonal level. Hence, further research is needed to explain the reason for the decline in the consolidated personal income tax rates over the observed period. Moreover, the findings cannot be transferred to other countries and regions due to the specific nature of the Swiss tax system.

Finally, our results weaken the argument that conventional spatial estimation methods are misleading in the analysis of tax competition. At least for the application at hand, the criticised SPML approach reveals insights that point into the same direction as the proposed IV method. While the ladder is superior from a methodological point of view, SPML estimation is more flexible and might be useful in cases where there exists no *experimentalist* approach as the one adopted by Parchet (2019a).

References

Gaure, S., 2013. lfe: Linear group fixed effects. The R Journal 5, 104–117. URL: `https://journal.r-project.org/archive/2013/RJ-2013-031/RJ-2013-031.pdf`. user documentation of the 'lfe' package.

Gibbons, S., Overman, H.G., 2012. Mostly pointless spatial econometrics? Journal of Regional Science 52, 172–191.

Kelejian, H.H., Prucha, I.R., 1998. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. The Journal of Real Estate Finance and Economics 17, 99–121.

Millo, G., Piras, G., et al., 2012. splm: Spatial panel data models in r. Journal of Statistical Software 47, 1–38.

Parchet, R., 2019a. Are local tax rates strategic complements or strategic substitutes? American Economic Journal: Economic Policy 11, 189–224.

Parchet, R., 2019b. Replication data for: Are local tax rates strategic complements or strategic substitutes? URL: `https://www.openicpsr.org/openicpsr/project/114646/version/V1/view`.

Appendix A. Additional tables

Table A.3: Summary statistics of the three samples used.

	Full sample		*Balanced sample*		*Border sample*	
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
$T_{ic,t}$	25.84	3.45	25.76	3.42	25.61	3.57
Population	2.85	10.28	2.89	10.40	2.61	4.88
% foreign	10.45	8.55	10.56	8.57	9.97	8.05
% young	19.95	3.83	19.77	3.77	20.34	3.48
% old	3.30	1.84	3.33	1.83	3.08	1.47
% secondary	31.04	11.26	30.35	10.85	32.37	10.63
% tertiary	56.57	14.39	57.83	13.88	54.87	13.01
Employment p.c.	0.26	0.20	0.25	0.21	0.24	0.17
% unemployed	2.19	1.96	2.32	1.98	2.07	1.72
% voting left	18.27	10.47	17.92	9.33	17.90	10.07
# cinemas	3.27	6.47	3.37	6.70	3.14	5.70
Number of observations	70,127		61,641		28,362	
Number of municipalities	2,428		2,283		1,047	
Number of years	30		27		29	

Table A.4: Estimation results of model (2) using different samples, weights matrices and estimation techniques. Standard errors are clustered by municipality and year.

	Dependent variable: $T_{ic,t}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Method	OLS	S2SLS(1)	OLS	S2SLS(1)	OLS	S2SLS(1)	OLS	S2SLS(1)
Sample	full sample		balanced sample		balanced sample		border sample	
$\boldsymbol{W}$	original $\boldsymbol{W}$		original $\boldsymbol{W}$		replicated $\boldsymbol{W}$		original $\boldsymbol{W}$	
$\boldsymbol{W}T_{ic,t}$	0.444***	0.613***	0.457***	0.676***	0.455***	0.684***	0.425***	0.662***
	(0.031)	(0.124)	(0.033)	(0.130)	(0.032)	(0.134)	(0.045)	(0.166)
Population	−0.039*	−0.039*	−0.037*	−0.039*	−0.037*	−0.039*	−0.071	−0.069
	(0.014)	(0.015)	(0.014)	(0.015)	(0.014)	(0.015)	(0.048)	(0.057)
% foreign	0.009*	0.009	0.008	0.008	0.008	0.008	0.024**	0.022*
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)	(0.009)
% young	0.012*	0.012*	0.014*	0.014*	0.014*	0.015*	0.008	0.009
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
% old	0.010	0.011	0.009	0.010	0.010	0.011	−0.022	−0.020
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.018)	(0.018)
% secondary	−0.023***	−0.021***	−0.024***	−0.021***	−0.024***	−0.021***	−0.025***	−0.022***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)
% tertiary	−0.021***	−0.019***	−0.020***	−0.018***	−0.020***	−0.018***	−0.017**	−0.015*
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.006)	(0.006)
Employment p.c.	−0.371**	−0.371**	−0.331*	−0.331*	−0.332*	−0.331*	−0.572	−0.632*
	(0.131)	(0.129)	(0.135)	(0.131)	(0.135)	(0.131)	(0.308)	(0.275)
% unemployed	0.015	0.015	0.016	0.016	0.016	0.016	0.025	0.026
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.016)	(0.016)
% voting left	−0.001	−0.002	−0.003	−0.003	−0.002	−0.003	−0.008*	−0.008*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
# cinemas	0.011*	0.012*	0.012*	0.012*	0.012*	0.012*	0.003	0.003
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.010)	(0.010)
Observations	70,127	70,127	61,641	61,641	61,641	61,641	28,362	28,362
Adjusted R^2	0.963	0.963	0.964	0.963	0.964	0.963	0.969	0.968
$\boldsymbol{WX}$	NO	NO	NO	NO	NO	NO	NO	NO

Note: *p<0.05; **p<0.01; ***p<0.001