Documento de Trabajo

Searching for thresholds in local corporate taxation: How do agglomeration economies affect?

JESÚS LÓPEZ-RODRÍGUEZ
DIEGO MARTÍNEZ-LÓPEZ
BRAIS POCIÑA-SÁNCHEZ

Documento de Trabajo 2025/02
Febrero de 2025

fedea

Searching for thresholds in local corporate taxation: How do agglomeration economies affect?*

Jesús López-Rodríguez[a,b,*], Diego Martínez-López[c,d,*], Brais Pociña-Sánchez[a,b,*]

[a] *Jean Monnet Group on Competition and Development (GCD), Universidade da Coruña and* [b] *Economics and Business Administration for Society (ECOBAS) at Campus de Vigo*

[c] *Pablo Olavide University and* [d] *FEDEA.*

Abstract

The foot-loose capital (FC) models predict that agglomeration forces create rents for the mobile factor (capital), which can be easily taxed, and thus higher equilibrium tax rates are expected. This paper uses a highly flexible econometric specification (P-Spline spatial autoregressive model, PS-SAR) to look at the relationship between tax rates and agglomeration economies in Spain over the period 2013-2020. Our results show the existence of a minimum level of agglomeration economies that are required to find taxable agglomeration rents. This outcome calls for a reassessment of the linear FC models to disentangle which mechanisms might lead to these phenomena.

Keywords: Geographical economics, taxable agglomeration rent, Spanish municipalities, semi-parametric spatial econometrics
JEL codes: H2, H3, C23, R12

1. Introduction

Geographical economics models, specifically the linear foot-loose capital (FC) models, are commonly used to support empirical research on the existence of taxable agglomeration rents (TARs) across local jurisdictions. These models predict that when regions are asymmetric in terms of immobile factor endowments (labor), agglomeration creates rents for the mobile factor (capital) that can be taxed, thus increasing equilibrium tax rates (Andersson and Forslid, 2003; Baldwin and Krugman, 2004; Ottaviano and van Ypersele, 2005). As a result, more agglomerated jurisdictions are better able to raise tax rates on the mobile factor without experiencing significant relocation, as businesses prefer to stay in agglomerated areas even with higher taxes, due to advantages such as labor pools, market size, and reduced transportation costs.

The theoretical results are obtained through the following mechanism: firms are sensitive to local tax decisions, and therefore, their location decisions are affected by changes in fiscal policies. If this happens and governments are aware of this causal relationship, they can then link their tax decisions to firms' location choices and, con-

*This version: February, 2025

*Corresponding authors:
Email addresses: jesus.lopez.rodriguez@udc.es (Jesús López-Rodríguez[a,b,*]), dmarlop1@upo.es (Diego Martínez-López[c,d,*]), brais.psanchez@udc.es (Brais Pociña-Sánchez[a,b,*])

sequently, to the observed levels of agglomeration. Following Brülhart et al. (2015), these two phases of the process must be empirically tested. The first necessary condition, the sensitivity of firms to tax rates, has been tested in various contexts such as Devereux et al. (2007), Jofre-Monseny and Solé-Ollé (2010) or Brülhart et al. (2012), demonstrating that firms located in agglomerated regions are less sensitive to tax rates than those in less agglomerated areas.

The second aspect, the hypothesis that governments are aware of and exploit this asymmetry in firms' location sensitivity—i.e., the existence of taxable agglomeration rents (TARs)—has been tested by several studies. Charlot and Paty (2007, 2010) demonstrate the existence of TARs in the French municipal system using an SAR-FE model. Jofre-Monseny (2013) also confirms the presence of TARs in Spanish municipalities through a fixed effects (FE) approach. Similarly, Koh et al. (2013) for Germany and Luthi and Schmidheiny (2013) for Switzerland also detect this effect in their respective studies. A common approach in this literature is the use of panel data with fixed effects, alone or combined with spatial dependence, which allows for controlling potential horizontal fiscal interactions across neighboring jurisdictions.

However, a key limitation of these models is their reliance by specifying a linear functional form between the dependent variable and its covariates. This assumption can be overly restrictive and may lead to biased estimates when the relationship between economic variables is more complex than a simple linear form[1]. This holds true for the relationship between agglomeration and tax rates, where non-linearities or threshold effects may be present, further complicating the dynamics and making linear assumptions inadequate for accurate estimation. Although the linear specification is often supported by underlying theoretical models[2], it imposes strong constraints on the empirical analysis, potentially overlooking non-linear dynamics present in fiscal competition.

While earlier contributions in the literature favored these linear models for their simplicity and interpretability, it has become increasingly evident that they may not adequately capture the complexity of real-world interactions. To address this shortcoming, recent research has increasingly turned to semiparametric frameworks, which offer greater flexibility by relaxing the linearity assumption. These models can capture non-linear relationships while accounting for spatial dependence and spatial heterogeneity. Notable examples include penalized-spline spatial Durbin models (SDM), such as the one introduced by Basile et al. (2014), which help identify the functional form between the response variable and its predictors[3]. These models are particularly useful in distinguishing between substantive spatial dependence—resulting from spatial spillovers—and spatial dependence that arises due to omitted variables. By incorporating non-linear dynamics, semiparametric methods provide a more nuanced

[1]See, for example, Burdekin et al. (2004), Christie (2012) or Sanso-Navarro and Vera-Cabello (2015), which are related to the detection and measurement of the effects of non-linearities on economic growth determinants.

[2]It is important to keep in mind that geographical economics models generate economies with significant non-linearities, including bifurcations, in the relationships between variables. These theoretical models support a linear relationship only when strong initial assumptions are imposed, aiming to simplify the models and derive closed-form solutions.

[3]For a complete review, see Mínguez et al. (2020).

understanding of the complexities of fiscal competition, especially when spatial effects play a significant role.

Building on this, we aim to overcome the limitation of assuming a linear specification when estimating the effects of agglomeration economies on tax-setting decisions. To achieve this, we estimate a spatial autoregressive semi-parametric (PS-SAR) model (Lee and Durban, 2011; Montero et al., 2012; Minguez et al., 2020) to determine whether non-linearities are present. The PS-SAR model is particularly effective in controlling for unobserved heterogeneity across both municipalities and time, accounting for the diverse socio-economic and political characteristics of local jurisdictions, as well as the varying sizes of neighboring areas.

We introduce an empirical methodology that provides estimates from the spatial autoregressive semi-parametric model to test whether the global linear assumption holds for the relationship between agglomeration and business tax rates. If this assumption does not hold, we investigate whether local linearity is compatible with the data—i.e., whether the relationship is linear within different segments where the slope varies significantly. If the local linearity assumption holds, we interpret the results in terms of slopes and thresholds. If it does not, we directly interpret the non-parametric function derived from the econometric approach.

In particular, identifying a threshold pattern in the detected non-linear relationship would suggest the existence of a minimum level of agglomeration economies required to generate *taxable agglomeration rents*. This would call for a reassessment of the predictions of taxation-linked geographical economics models, helping to clarify which mechanisms within these models may lead to such phenomena.

The empirical model has been estimated for a panel of Spanish municipalities for the period 2013-2020. The results of our estimations show indications of non-linearities associated with the impact of agglomeration economies on the equilibrium tax rates. These non-linearities confirm the existence of threshold effects, meaning that there is a significant change in the relationship between local business tax rates and agglomeration economies beyond a certain threshold of agglomeration. Specifically, we find that for agglomeration levels in the top 20th-30th percentile of the distribution, the relationship between corporate taxation and agglomeration economies exhibits *taxable agglomeration rents*. Below this threshold, this relationship is extremely weak, and we can safely rule out the existence of *taxable agglomeration rents*.

Therefore, our results lead to the conclusion that the *taxable agglomeration rents* phenomenon predicted by the geographical economics literature only operates at high levels of agglomeration, where economies of scale strongly come into play. These results demonstrate that imposing a linear relationship between our interest variables would cause an overestimation of the agglomeration effect on equilibrium tax rates for the majority of municipalities. To the best of our knowledge, this is the first paper documenting that *taxable agglomeration rents* are not an *always and everywhere* rule in local tax setting decisions, as linear FC models predict. We have also controlled for the potential endogeneity of our agglomeration economies variable.

The last part of the paper, taking advantage of our highly non-restrictive econometric approach and based on a ”let the data talk” approach, re-explores the link between agglomeration economies and tax-setting decisions from a theoretical perspective. Under the standard assumptions of these models, such as imperfect competition, increasing returns to scale, and trade costs, we guess some potential explanations on how to bridge the gap between our empirical results and the mainstream of the theoretical literature.

The rest of the paper is organized as follows. In Section 2, we present the PS-SAR model. Section 3 briefly introduces the Spanish institutional context and describes the data and variables used in our estimations. Section 4 deals with the empirical strategy we follow to test for taxable agglomeration rents and the identification of threshold effects. In Section 5, the empirical results are explained. In Section 6, we present the theoretical discussion, offering a rationale for the empirical results we have achieved. Finally, Section 7 concludes.

2. The PS-SAR approach to testing for taxable agglomeration rents

2.1. Spatio-temporal semiparametric autoregressive models: The PS-SAR model

A PS-SAR model can be specified as follows:

$$y_{it} = \underbrace{\rho \mathbf{W} y_{jt}}_{A} + \underbrace{\sum_{\delta=1}^{k} m_\delta(x_{\delta,it})}_{B} + \mathbf{X}'_{it}\beta + \underbrace{f(s_{1i}, s_{2i}, \tau_t)}_{C} + \epsilon_{it} \tag{1}$$

where the term A captures the spillover effect of a SAR specification, i.e., it represents a weighted average of the dependent variable y, let's say local tax rates, across the "j" locations which are neighbours to location "i". This term is made up of an autoregressive parameter, ρ, the spatial weights matrix, W, which defines the neighbourhood criteria across locations and the values of the dependent variable in the "j" locations. The term B captures the non-parametric part of the model, where $m_\delta(\cdot)$ is a non-linear function of the δ-th variable $x_{\delta,it}$. The term C captures non-observable heterogeneity across locations (within-location variation) taking the form of a global spatio-temporal trend in the model, being (s_{1i}, s_{2i}) the spatial coordinates (latitude and longitude) of the centroid in the location i and the temporal dimension τ_t. Coordinates and the temporal dimension are related through an unknown function $f(\cdot)$ with the dependent variable.

Finally, the β term captures the effects of other explanatory variables ($\mathbf{X}_{it}$) interacting in a linear way with the dependent variable, and ϵ_{it} is the disturbance of the model which is assumed to be a white noise. Due to computational costs, estimating the spatio-temporal term (estimate the term C in (1)) directly is not advisable for large datasets. To address this issue, Lee and Durban (2011) developed an ANOVA decomposition approach guided by the following structure expressed in vector notation:

$$\begin{aligned} f(\boldsymbol{s_1}, \boldsymbol{s_2}, \boldsymbol{\tau}) =& f_1(\boldsymbol{s_1}) + f_2(\boldsymbol{s_2}) + f_t(\boldsymbol{\tau}) + f_{1,2}(\boldsymbol{s_1}, \boldsymbol{s_2}) + f_{1,t}(\boldsymbol{s_1}, \boldsymbol{\tau}) \\ &+ f_{2,t}(\boldsymbol{s_2}, \boldsymbol{\tau}) + f_{1,2,t}(\boldsymbol{s_1}, \boldsymbol{s_2}, \boldsymbol{\tau}) \end{aligned} \tag{2}$$

Non-parametric estimations of the PS-SAR model are carried out from the methodology of penalized splines (P-splines) developed by Eilers and Marx (1996) and using the statistical package *pspatreg* developed by Minguez et al. (2022) with the programming language R.

2.2. P-splines methodology

The P-splines methodology, developed by Eilers and Marx (1996), is a modification of the B-splines technique which consists of introducing a penalization in the objective function to be minimized. The penalization term allows to significantly improve the goodness of fit properties.

A B-spline is a function made up by joining polynomials at certain levels of the ”x” variable which are known as *knots* (De Boor, 1978). This type of functions is widely used as the base when adjusting non-parametric regressions. The base is determined by the way in which the knots are distributed over the range of the “x” variable and by the degree of the polynomial. It is quite common the use of polynomials of degree three for B-splines, however, there is some discrepancy in how to establish the optimal number of knots (Ruppert, 2002; Eilers and Marx, 1996; Minguez et al., 2020).

The adjustment of a variable through the B-splines methodology is then determined according to the following expression:

$$\hat{y}_i = \sum_{j=1}^{n} \hat{a}_j B_j(x_j) \tag{3}$$

where n is the number of knots, B_j is the base of order p splines (generally $p = 3$) at the x_j point and $\hat{a}_j$ is the vector of coefficients obtained by minimising the following expression:

$$S = \sum_{i=1}^{m} \left(y_i - \sum_{j=1}^{n} a_j B_j(x_j) \right)^2 \tag{4}$$

The P-splines technique, however, modifies the previous function by introducing a penalization term, usually by second degree differences in adjacent B-splines coefficients. In particular:

$$S^p = \sum_{i=1}^{m} \left(y_i - \sum_{j=1}^{n} a_j B_j(x_j) \right)^2 + \lambda \sum_{j=k+1}^{n} (\Delta^k a_j)^2 \tag{5}$$

The introduction of this penalization term offers several advantages over the B-splines method (Eilers and Marx, 1996; Eilers et al., 2015), including the absence of frontier effects, the ability to fit polynomial data exactly, robustness to the selection of knots beyond a certain value, lower computational costs compared to other penalized models such as O’Sullivan (1986), and the re-parametrization of the model as a ”mixed model,” which enables simultaneous estimation and optimization of all model parameters.

3. The Spanish institutional context, variables and data

3.1. The Spanish institutional context

The Spanish local institutional context is characterized by three tiers of overlapping governments: municipalities, provinces and Autonomous Communities. The lowest tier of the territorial organization is made up of 8131 municipalities (local jurisdictions). These municipalities are aggregated into 50 provinces (NUTS[4] III), excluding the autonomous cities of Ceuta and Melilla, and 17 Autonomous Communities (NUTS II). The government and administration system of the municipalities corresponds to the Town Councils, consisting of Mayors and Councillors. The government system and autonomous administration of the provinces is entrusted to Provincial Councils, known as *"Diputaciones"*.

Municipalities are responsible for local urban services (street lighting, waste collection, supply of drinking water), building and maintaining nursery and primary schools and sport facilities, municipal roads and urban public transport. The *"Diputaciones"* offer legal, economic and technical assistance and cooperation to municipalities, especially those with less economic capacity and management (waste management, fire prevention and extinction, maintenance and cleaning of medical offices, maintenance of provincial roads).

3.2. Data and variables

The empirical analysis is carried out by building a dataset with 7083 Spanish municipalities[5]. The data have been obtained from the Spanish National Statistics Institute (INE), the Ministry of Finance, the Ministry of Labour and Social Security, and OpenStreetMap for distance calculations. We have pooled municipality for a 8 year-period, starting in 2013 and covering therefore the period (2013-2020), yielding a total of 56,664 observations.

3.2.1. Business tax

There are three mandatory taxes and two taxes of voluntary establishment by municipalities. The compulsory levying taxes are the real estate or property tax (*Impuesto sobre Bienes Inmuebles*), which is the most important in terms of tax revenue, the local business tax (*Impuesto sobre Actividades Económicas*), which is the main local tax burden borne by the business sector and the vehicles tax (*Impuesto de Vehículos de Tracción Mecánica*). The other two voluntary taxes are those applied to the increase in the value of urban land (Impuesto sobre la plusvalía) and to request a license to build or repair a premise (*Impuesto de Instalaciones, Construcciones y Obras*).

The local business tax is a presumption tax computed from different indicators of economic activity. The tax base is determined by national tax laws and is meant to approximate a share in a firm's profits. This tax base is then weighted by a municipal-specific augmentative coefficient (*coeficiente de localización*), t_{it}, which applies to all

[4]The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the EU and the UK.

[5]We remove from our sample those municipalities that do not belong to the so-called '*Regimen Comun*', are not located in the Spanish mainland and do not have enough data to keep a balanced panel.

establishments in each municipality. This municipal-specific weighted coefficient is applied based on the establishments' physical setting within the municipality in relation to the category of the street in which is located and cannot be lower than 0.4 and higher than 3.8. We refer to this municipal coefficient as the local business tax rate. Although there is an important degree of autonomy in setting this tax rate, the size of the municipal population establishes an upper bound for it.

This local business tax is estimated to have collected more tha 1,750 millions of euros in 2024. As mandatory tax, its distribution over the territory is quite homogeneous, obviously controlling for the size of municipalities; about 28 per cent of the total revenue is obtained in juridictions with less than 20,000 inhabitants. Since its reform in 2004, only companies with turnovers equal or higher than 1 million of euros become taxpayers. The White Book on the reform of local public finances ordered by the Spanish Ministry of Finance, published in 2017, suggests a number of technical improvements in its design aimed at recovering the relevance of such a tax in the map of local taxes in Spain (Spanish Institute of Fiscal Studies, 2018, p.238-242).

3.2.2. Measuring agglomeration economies

To capture the size of agglomeration, we use the measure proposed by Harris (1954) called Market Potential. This measure involves a sum of a certain economic activity indicator weighted by the inverse of the distance between a region and the rest of the regions in a specific geographical space. Specifically, we use the population measure as an economic activity indicator, minutes by road as the distance measure between each municipality (measured from centroid to centroid), and the Spanish peninsular territory as the reference geographical space.

This variable measures the potential demand of a region (in our case, a municipality) and is usually used in the literature that test for taxable agglomeration rents see, for instance, Charlot and Paty (2007, 2010). The expression used to calculate municipality's market potential is given by expression (6):

$$\mathrm{MP}_{i,t} = \frac{\mathrm{Pop}_{i,t}}{d_{i,i}} + \sum_{j \neq i}^{N} \frac{\mathrm{Pop}_{j,t}}{d_{i,j}} \tag{6}$$

where $MP_{i,t}$ represents the total market potential of municipality "i" in year t; $Pop_{i,t}$ and $Pop_{j,t}$ represent the population of the municipality "i" and "j" in year t respectively; $d_{i,i}$ is the internal distance within each municipality[6] "i" and $d_{i,j}$ is the distance between municipality "i" and "j" (measured both in minutes of travel and in kilometers).

3.2.3. Control variables

Besides controlling for fiscal interactions, we have added to set of additional variables to control for levels of public spending and political orientation. In particular,

[6] The internal distance within each municipality (d_{ii}) is estimated using a Monte-Carlo method that consists of randomly selecting N pairs of points that are within the geographical limits of each municipality. We calculate the distance for each pair of points and take the mean of all the values to estimate the intra-municipality mean distance.

to proxy public spending we have used the *percentage of population below 16 years of age*, the *percentage of population above 65 years of age*, the *percentage of immigrants in the municipality* and the *municipal unemployment rate* computed from data of our national unemployment public services (SEPE) jointly with data on affiliations taken from our social security records. The political dimension has been control for by two dummmies: i) *left*, that takes the value 1 for left-wind parties (PSOE-member of the European S&D- and all political parties that in the political spectrum are to the left of PSOE), ii) *national*, that takes the value 1 for all political parties that in the municipal elections have representation in all the Spanish territory. Hence, nationalist, regionalist as well as independentist political parties take the value 0 for this variable.

3.3. Summary statistics

Table 1 provides some descriptive statistics for the main variables used in our econometric estimations[7].

Table 1: Summary statistics

	Mean	St. Dev.	Min	Max
Business tax rate proxy (IAE)	1.20	0.55	0.40	3.80
Market Potential (distance in minutes)	163,206	38,401	77,658	493,842
Market Potential (distance in km)	176,104	57,021	98,730	789,588
Provincial surchage	31.50	9.56	0.00	41.00
Unemployment Rate	22.32	13.27	0.00	100.00
% Population under 16 years	10.55	5.65	0.00	29.22
% Population over 65 years	29.34	11.30	0.00	90.00
% Population inmigrant	6.58	7.40	0.00	90.71

4. Empirical Strategy

4.1. Baseline specification

The econometric specification usually employed to assess the existence of *taxable agglomeration rents* is (7):

$$t_{it} = \rho W t_{jt} + \beta_1 \mathrm{MP}_{it} + \beta_2 t^p_{it} + X'_{it}\gamma + \alpha_i + \delta_t + \epsilon_{it} \tag{7}$$

where the term that captures agglomeration is MP_{it} and is related linearly with the dependent variable, the tax rate set by the corresponding jurisdiction. In the econometric specification two additional terms are included to control for horizontal and vertical fiscal interactions. Specifically, we control for horizontal fiscal interactions through the term $\rho W t_{jt}$, where W is the weights matrix, and for vertical fiscal interactions through the variable t^p_{it}, which is the business tax rate set by the province

[7]The zero percent minimal values of the control variables reflected in the table are explained by the inclusion of the majority of Spanish municipalities in the sample. As a result, municipalities with very small population levels are included, leading to these extreme values.

to which each municipality belongs also known as *recargo provincial del IAE*. Additionally, matrix X collects a set of variables related to socio-economic and political dimensions that influence the determination of equilibrium business tax rates[8].

With the aim of relaxing the linearity assumption in the relationship between agglomeration and business tax rates on the previous specification, we propose to modify the SAR specification of the model (7) to a PS-SAR model (8), which allows for the introduction of semi-parametric terms.

$$t_{it} = \rho W t_{jt} + m(\mathrm{MP}_{it}) + \beta t^{p}_{it} + X'_{it}\gamma + f(\boldsymbol{s_1}, \boldsymbol{s_2}, \boldsymbol{\tau}) + \epsilon_{it} \tag{8}$$

In this specification we regress Spanish municipal business tax rates t_{it} against municipal market potential MP_{it} which is added to the model through an unknown $m(\cdot)$ function. We retain all existing control terms in the baseline specification (7). All terms but market potential are added to the model in a linear way. The spatial and temporal fixed effects of the specification (7) are introduced in this expression through the multivariate function f, which captures the space-time trend. Finally, ϵ_{it} is the disturbance term associated with the municipality 'i' in year 't'.

For robustness purposes, we have re-estimated the previous specification (8) changing the criterion of neighborhood used to define the spatial weights matrix and the proxy of agglomeration, as is described in the next subsections.

4.2. Treatment of endogeneity

4.2.1. Weighting matrix

A first source of endogeneity in the proposed model is the correlation between the spatial autoregressive term (Wy) and the disturbance term (ϵ). To control the bias caused by this endogeneity, following Basile et al. (2014), we employ the control function method, using exogenous variables and several orders of their spatial lags as instruments. These instruments are introduced through semiparametric univariate functions following the expression:

$$Wy = \theta_0 + \sum_{j=1}^{N} g_j(Z_j) + \nu \tag{9}$$

where θ_0 is the intercept, the function g_j is the j-th unknown univariate function, and the variable Z_j is the j-th instrumental variable, which are the covariates of equation (8) along with their spatial lags[9]. Finally, ν is the Gaussian error term.

After estimating the proposed specification, the estimated error term ($\hat{\nu}$) is used as an additional explanatory variable in the baseline equation (8), introduced through an unknown univariate function (h).

4.2.2. Market potential

The variable we use to measure agglomeration economies may suffer from endogeneity problems mainly due to reverse causality. On the one hand tax rates depend

[8]See section 3.2.3 for a complete description of the added control variables.

[9]The spatial lags have been constructed using the 5-nearest neighbors criterion.

of the levels of agglomeration economies (our main hypothesis) but it is also possible that the size of agglomeration economies could be explained by the tax rates set by different jurisdictions. For instance low tax rates could be a driver for more capital to be allocated to a region and therefore this leads to a reinforcement of agglomeration economies for the region and, the other way around, a high taxation deters more capital to come into a region and decreases the levels of agglomeration economies of the region.

To control for this potential endogeneity issue, we use the instrumental variables method. Market potential has been instrumented using as instrument historical values for market potential (a lag of our measure of agglomeration). Through these temporal lags we break the problem of reverse causality. Tax rates at moment t can be explained by the size of the agglomeration economies at time $t-n$ (i.e agglomeration economies tend to be reinforced as the time goes by but without big changes in the patterns, historical causes, nature of the economic geography, agglomeration trends) while the level of agglomeration economies at time $t-n$ cannot be explained by the tax rate that has been set at time t[10].

4.3. Testing linearity

The main goal we pursue in this paper is to assess if the relationship between local corporate taxation and agglomeration economies shows non-linearities. In particular, by softening the hypothesis of a linear relationship in our econometric specification we want to reexamine if jurisdictions may have different behavioral patterns in terms of business tax setting decisions for different levels of agglomeration.

The methodology we have followed to examine non-linearities in the relationship between business taxation and agglomeration economies takes the following steps: i) First, we estimate the PS-SAR model proposed in specification (8); ii) Second, we check for linearity by regressing the estimated values of the non-parametric part of specification (8), $\hat{m}(\cdot)$, against the agglomeration variable using weighted OLS[11]. We use a RESET test to determine if a linear relationship between agglomeration and business tax rates holds, iii) Finally, if the hypothesis of linearity (global) is rejected, we estimate a piecewise regression[12] to determine if we can still find local tranches of linearity despite the rejection of a global linear relationship.

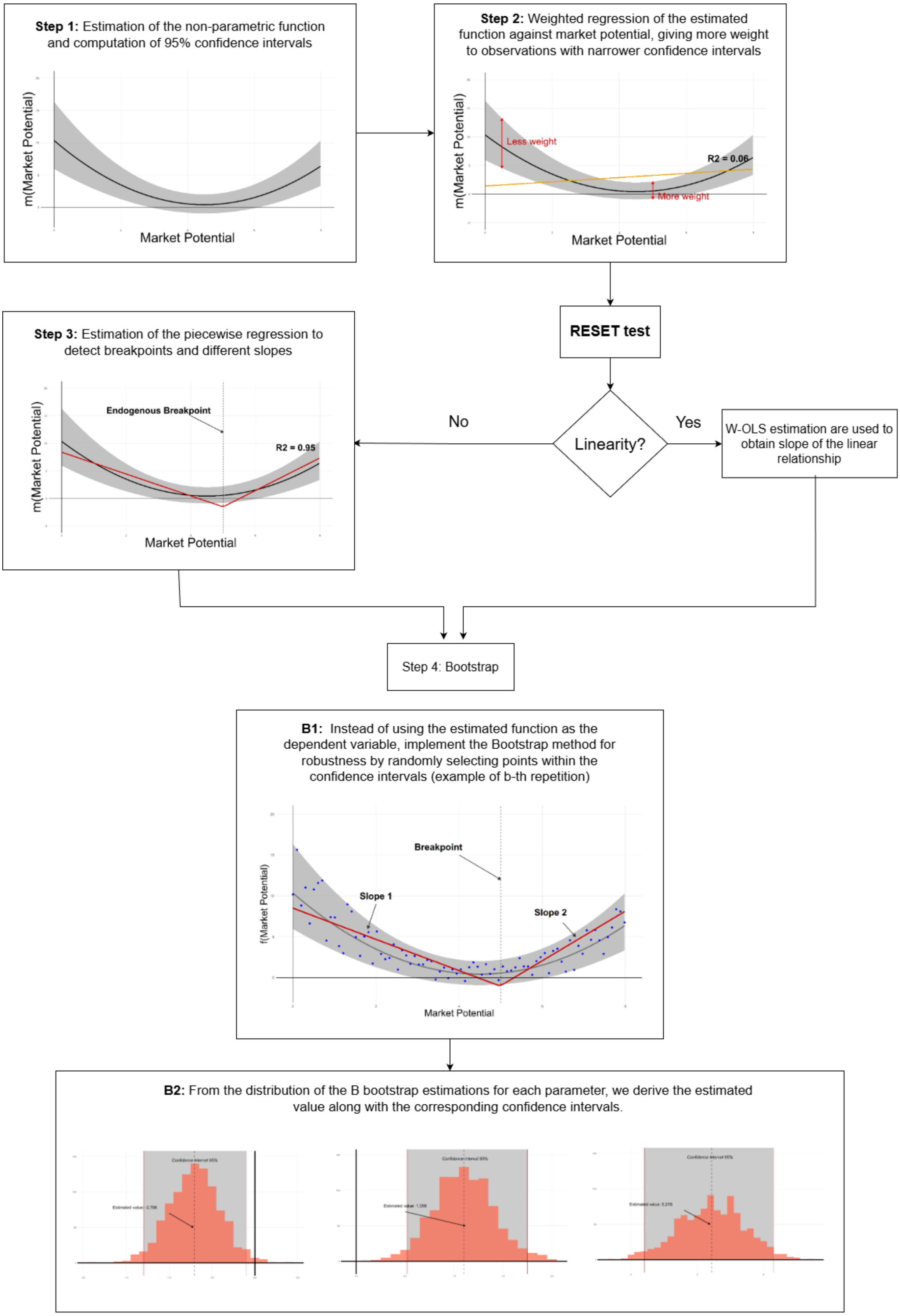
To improve the precision in this step, we propose a bootstrap approach. For a b-th repetition we have used a sample of points randomly selected within the confidence interval obtained from the PS-SAR estimation of the m function and we obtain the b-th piecewise regression results. We have repeated this process B times. The mean values of the breakpoints and the slopes estimated are taken as the estimated values and the confidence intervals are derived from the percentiles of the empirical distribution. For a better understanding of the complete methodology, Figure 1 is presented.

[10] In this case we use a ten-year lag of market potential.

[11] Weights are assigned based on the inverse of the square of the standard errors for each point in the non-parametric estimation.

[12] Piecewise regressions are aimed to estimate models with local linear relationships without imposing an exogenous number of thresholds. For more details, see Muggeo (2003).

Figure 1: Empirical approach to deal with non-linearities



5. Results

Table 2 presents the results from the standard model typically estimated in the related empirical literature (7), as well as the estimation of our baseline model (8) and the alternative specification proposed to test the robustness of our baseline results. Specifically, column 1 presents the estimation of the SAR model (7), while column 2 presents the estimation of our semiparametric baseline model (8). Columns 3 and 4 present control estimations, with modifications in the measure of agglomeration (distance measured in kilometers) and in the neighborhood criterion used to construct the spatial weights matrix (10-nearest neighbors), respectively. Finally, columns 5 and 6 present the estimations that control for potential endogeneity in both the spatial autoregressive term and the measure of agglomeration. [13]

We break down the table in three blocks. The first block shows the results of the parametric part of the model, with the estimated coefficients for each variable, statistical significance and standard errors; the second block presents the main metrics associated with the non-parametric estimation of the agglomeration economies variable (market potential) jointly with the battery of statistics and results related to the empirical methodology explained in subsection 4.3. For each estimation, the R^2 of the W-OLS regression (step 2 in Figure 1) is presented, along with the RESET test. This block also includes the estimated breakpoint (Threshold)[14] and the R^2 of the piecewise regression along with the two slope estimates and the statistical significance of them. The last block of the table includes a series of goodness-of-fit statistics as well as information related to the configuration of the model used.

The results from Table 2 are complemented by figure 2 that represents $\hat{m}(\cdot)$ function, which relates the level of agglomeration to business tax rates for each proposed specification. This figure shows the estimated function for the baseline model (column 2) along with the additional information provided in the non-parametric section of the table. The non-parametric representations for the models 3-6 are presented in the appendix[15].

In figure 2 the black line represents the estimation of the function $m(\text{MP})$, and the shaded gray area represents the 95% confidence interval for this estimation (step 1 of the empirical methodology). The orange line represents the approximation of the function using a W-OLS regression (step 2 of the methodology), while the red dashed line represents the approximation using the proposed method based on piecewise regressions (step 3 of the empirical methodology). Additionally, the shaded blue area shows the estimated market potential density function, and the vertical dotted line

[13] We employ the 10-nearest neighbor matrix to introduce spatial lags in the first stage of the control function estimation to control for endogeneity in the weights matrix (see section 4.2.1 for a detailed explanation), and a temporal lag of 10 years as an instrument to control for endogeneity in market potential.

[14] For ease of comparison, the breakpoint is expressed in terms of the percentile associated with the market potential distribution. In all models, the hypothesis of linearity is rejected, and the number of segments determined endogenously is equal to two.

[15] It is important to note that the orange line does not represent the slope estimated for the model that assumes linearity in the relationship between market potential and business tax rates. Instead, the orange line refers to the relationship between market potential and the non-parametric function estimated for the relationship between market potential and business tax rates.

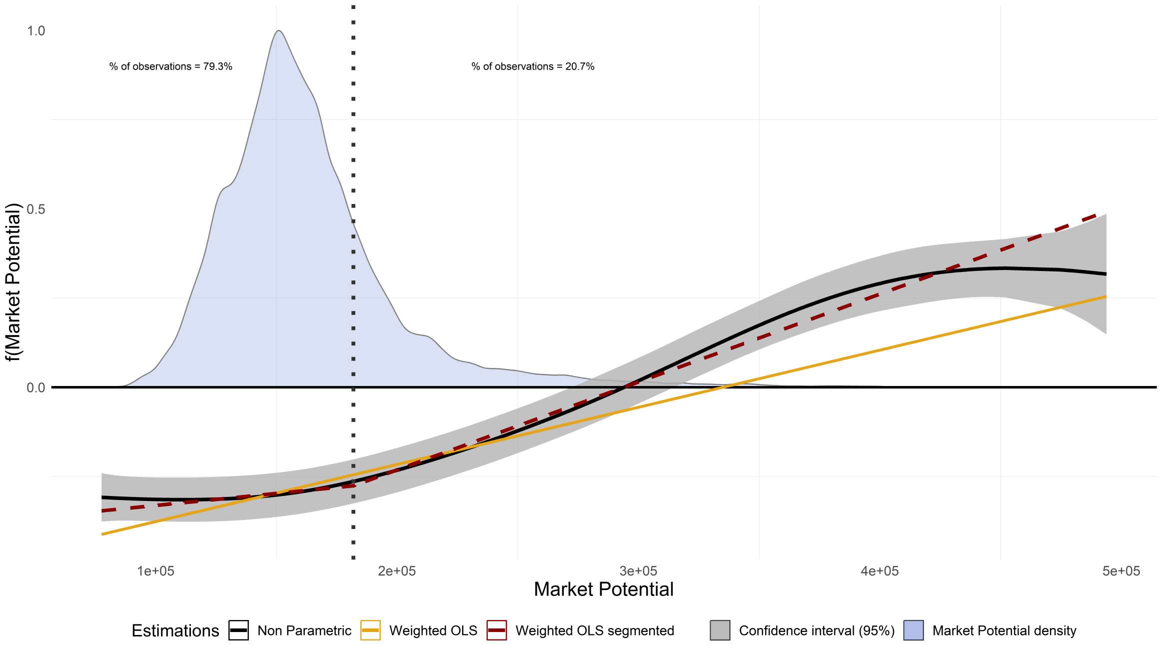
Table 2: Estimates of PS-SAR model (8). Dependent variable: Local business tax rate.

	SAR	Baseline	Control estimations		Endogeneity	
	(1)	(2)	(3)	(4)	(5)	(6)
Parametric						
ρ	0.026***	0.001	0.002	0.001	0.001	0.001
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Market Potential	3.144***	-	-	-	-	-
	(0.000)					
Provincial surcharge	−0.023**	0.041***	0.041***	0.037***	0.041***	0.041***
	(0.009)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Unemployment(%)	−0.000	−0.001***	−0.001**	−0.001***	−0.001***	−0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Pop_{<16}(\%)$	0.001***	0.009***	0.009***	0.008***	0.009***	0.009***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Pop_{>65}(\%)$	0.001***	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Pop_{inm}(\%)$	0.000	0.000	0.000***	0.001***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Left party	0.002	0.048***	0.048***	0.047***	0.048***	0.048***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
National party	−0.005**	−0.093***	−0.093***	−0.088***	−0.093***	−0.093***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Non parametric						
EDF	-	4.971	4.971	5.011	4.963	4.971
R^2 (W-OLS)	-	0.877	0.871	0.829	0.914	0.871
RESET test	-	11870***	11870***	2019***	12593***	11870***
Threshold (percentil-th)	-	79.30	79.30	75.05	79.53	79.42
R^2 (piecewise OLS)	-	0.991	0.991	0.985	0.994	0.991
Slope 1	-	0.673	0.676	0.093	0.915*	0.673
Slope 2	-	2.464***	2.465***	1.920***	2.432***	2.466***
Estimation method	SAR-FE	PS-SAR	PS-SAR	PS-SAR	PS-SAR-CF	PS-SAR-IV
Endogeneity control	-	-	-	-	Wt_{jt}	MP_{it}
W (k-nearest neighbors)	5	5	10	5	5	5
σ	0.417	0.350	0.350	0.314	0.329	0.350
AIC	-68049	-76830	-76830	-78206	-76849	-76830

Notes: *, ** and *** denote statistical significance at the 10, 5 and 1 percent level, respectively. Standard errors in parentheses. For ease of interpretation of the results, the estimated coefficients for Market Potential (column 1) and for the different slopes of Market Potential (columns 2-6) have been multiplied by a factor of 10^6.

indicates the threshold level.

Figure 2: Non parametric detection of the threshold. Baseline estimation



Analyzing the results from Table 2 and Figure 2 in tandem, we observe that using a SAR specification that assumes linearity in the relationship between the level of agglomeration and tax rates leads to an overestimation of the effect of market potential across all municipalities (see market potential coefficient in column 1 vis-à-vis coefficients for slope 1 and slope 2 in columns 2-6).

Our results allow us to identify two groups of municipalities based on how fiscal decisions interact with agglomeration. The first group, which takes in the vast majority of municipalities feature market potential values which are below the 75th-80th percentile. In this group a positive relationship between agglomeration and tax rates is detected only when controlling for the endogeneity of the spatial weights matrix, and this relationship is significant only at the 10% significance level.

The second group takes in municipalities in the upper quintile of market potential. This group shows a significantly stronger association between agglomeration and tax rates compared to the first group, being this relationship statistically significant at all conventional significance levels. These results imply that the relationship between tax rates and the level of agglomeration is clearly non-linear at the global level, but linear at the local level (within each tranche).

Additionally, our results imply that using a model that assumes global linearity in the relationship leads to an overestimation of the effect for all municipalities and lead to the existence of TARs across the entire territory. However, controlling for non-linearities we have a more nuance approach to the relationship, being the TARs effect only statistically significant for a small group of municipalities (the most agglomerated ones) while for the rest the relationship does not hold[16]. Moreover, assuming global

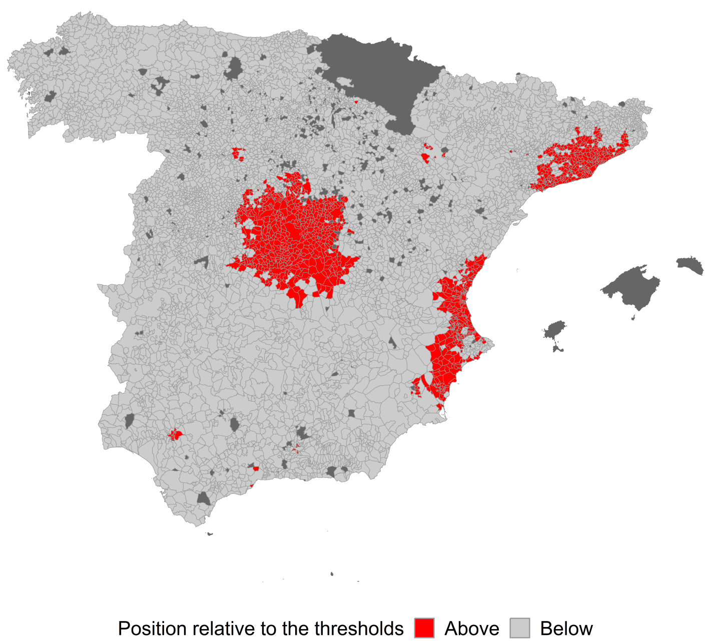
[16]We have only detected a significant effect for the less agglomerated group in column 5, but it is

linearity leads to an increase in the estimation error and a higher AIC statistic, meaning it also worsens the model fit.

Figure 3 groups municipalities according to the thresholds found in our econometric estimates for the relationship between local corporate taxation and agglomeration economies relationship. Two groups can be observed: a) A group labelled as 'Below-threshold' including municipalities with a level of agglomeration that is below the critical threshold found in any of the estimated models (light gray colour); and b) Another group labelled as 'Above-Threshold' that with those municipalities with levels of agglomeration that are above the threshold found in the estimated models (red colour)[17].

Therefore, this map allows us to detect those municipalities for which a *taxable agglomeration rent* phenomenon is intense in comparison with the rest of Spanish municipalities for which TARs effects are negligible or minimal. The 'Above-Threshold' group is composed by municipalities located in the surroundings of Madrid, Barcelona and Valencia, as well as other large, more isolated municipalities such as Zaragoza and Seville. These results are in line with the hypothesis that a *taxable agglomeration rent* phenomenon is present when a critical threshold for agglomeration is overcome.

Figure 3: Intensity in the TARs effect



The control variables related to socioeconomic or political factors are in line with the theoretical expectations and the standard results obtained in the related literature in all specifications. Specifically, a higher percentage of non-working population is associated with higher tax rates, consistent with the idea that these population group

only significant at a 10% and the slope is around 2.6 times lower than for the more agglomerated group.

[17]It is important to bear in mind that regardless of the specification estimated, the clustering of municipalities into these two groups is unaffected.

drains more public social spending (e.g., health, education, childcare, social assistance, etc.) compared to others.

On the other hand, a higher unemployment rate leads to lower business tax rates. This can be interpreted as a tool aimed at reducing costs for businesses, enabling them to absorb higher personnel expenses and, in turn, reduce unemployment[18]. Regarding the political dimension, municipalities governed by left-wing parties tend to set higher business tax rates than those governed by other parties; this is alligned with their leftist ideology, more oriented towards social spending programs.

Finally, municipalities governed by regional parties tend to set higher tax rates than those led by parties with nationwide representation. A potential explanation for this could be that regional parties need to differentiate themselves from the main parties, and a way of doing so might be to increase public spending, thereby requiring more resources (higher tax rates).

In relation to the fiscal interactions, relaxing the assumption of linearity (through a PS-SAR model) in the relationship between market potential and business tax rates leads to changes in the sign of the vertical fiscal interaction measured by provincial surcharges on local taxes. The positive and significant signs of the coefficients in the PS-SAR model specifications reveal indeed a strategic complementarity between different levels of government. Simultaneously, the horizontal interaction captured by ρ loses its statistical significance.

In sum, our empirical findings are the following: i) the existence of TARs is not observed for all municipalities, but only for the larger ones; ii) the effect of agglomeration on business tax rates is not globally linear, but locally linear and iii) assuming global linearity leads to an overestimation of the effect for all municipalities and a poorer fit accuracy. In the following section, we will discuss the implications of these empirical results using the mainstream theoretical model proposed in the related literature, aiming to bridge the gap between theoretical predictions and our empirical findings.

6. Sketching possible explanations

The analysis to be carried out next is centered in the identification of the mechanisms that might lead to a change in the regime guiding the relatioship between tax rates and agglomeration economies. We do not pursue to offer a closed explanation of the issues at play. Rather, we just aim to guess some potential explanations motivating further research. Recall that our central results in Figure 2 show two stretches in such a empirical relationship: one with null slope (no effect of agglomeration on taxation) until certain threshold and, from it onwards, other reflecting a positive impact of agglomeration on the local tax rates.

With this aim, we use the canonical model initially described by Ottaviano (2001) and further developed by Baldwin and Krugman (2004), Ottaviano and van Ypersele (2005) and Charlot and Paty (2007), among others. We just summarize its main

[18]It is also important to note that the funding for unemployment benefits in Spain does not depend on the municipality, but on the central government.

preliminary features, without entering into its technical details, because this theoretical framework is well-known enough to repeat here its standard theoretical developments.

We consider an economy with two local governments (North, N, and South, S), two production factors (labour, L and private capital, K) and two private sectors (agricultural and manufacturing sectors). The labour is assumed to be immobile across jurisdictions whereas the services of the capital are allowed to be provided in a jurisdiction different to where the owner of such capital is located. The agricultural sector employs only labour to produce a freely tradable good under constant return to scale (CRS). The manufacturing sector uses labour and capital to produce $n_N + n_S$ non-freely tradable i industrial varieties of goods under monopolistic competition and increasing returns to scale (IRS); each industrial firm is assumed to require one unit of K as a fixed factor to produce one variety. Shipping one unit of industrial variety of good i costs τ units of the agricultural good, which is used as *numeraire*, to take the trade costs into account.

The representative consumer, with a fixed endowment of production factors and the agricultural good, maximizes a utility function with love-for-variety. Given the mobility of private capital across jurisdictions, we need to distinguish between the share of capital owned by the residents in the jurisdiction N, denoted by σ and that we shall use as a measure of agglomeration, and the share of capital employed in the region N, named as γ. From the consumer optimization problem we obtain the following linear demand for the quantities of a variety i:

$$q(i) = a - bp(i) + c \int_{j=0}^{n_N+n_S} [p(j) - p(i)]dj \tag{10}$$

In this demand function $p(i)$ and $p(j)$ are respectively the prices of variety i and j; a, b and c are combinations of parameters of the utility function α, β and ν[19], particularly $a \equiv \frac{\alpha}{\beta+(N-1)\nu}$, $b \equiv \frac{1}{\beta+(N-1)\nu}$ and $c \equiv \frac{\nu}{(\beta-\nu)[\beta+(N-1)\nu]}$, and n_N+n_S are the total number of varieties made in jurisdictions N and S, respectively. The consumption of agricultural good is determined as a residual.

Whereas in the agricultural sector, with CRS and perfect competition, the maximization of profits leads to equalize the price to the marginal cost of production, in the industrial sector things are quite different. Under monopolistic competition, profit maximization involves pricing strategies related the trade costs and the intensity of competition across the firms in particular markets.

Moreover, there will be two-way trade between region N and region S if a critical threshold for trade costs is set up according to:

$$\tau < \tau_{trade} = \frac{2a}{2b + cN} \tag{11}$$

In the model, in the long-run, private capital moves freely across the two jurisdictions searching for the highest rewards. Considering that the capital employed in each

[19] α refers to the intensity of preferences for the manufacturing good, β to the love-of-variety and ν to the degree of substitutabilty between different varieties, with $\beta > \nu > 0$.

region is given by the parameter $\gamma \in [0,1]$, the arbitrage condition establishes that in equilibrium capital rewards r must be the same in the two jurisdictions. Therefore, the following condition must hold:

$$r_N(\gamma) = r_S(\gamma) \tag{12}$$

Solving (12) for γ we obtain γ^M, which is the closed-form solution for the spatial distribution of industry across jurisdictions and depending, among other things, on the trade cost τ and private capital endowment K.

We consider now two local governments maximizing a social welfare function in a Nash-type game. Under such environment, both local governments simultaneously choose first their tax rates, taking the decision of the other as given. The consumers and firms, whose behaviour has been briefly drawn above, maximize then their objective functions given the tax rates established by both local governments. The main modification in the framework sketched above lies in the arbitrage condition (12), which becomes:

$$r_N(\gamma) - t_N = r_S(\gamma) - t_S = \rho \tag{13}$$

where t_N and t_S are the capital tax rates of local government N and S, respectively. Solving implicitly the new arbitrage condition (13) for γ we arrive at:

$$\gamma\left(t_N, t_S\right) = \gamma^M - 2\frac{t_N - t_S}{\tau^2}\frac{(2b + cK)}{cKL\left(b + cK\right)} \tag{14}$$

This function $\gamma\left(t_N, t_S\right)$ guides the distribution of industry across the territory but taking now into account the local tax rates set by the different governments N and S.

We start by maximizing the social welfare function of the jurisdiction N given by:

$$W_N = G_N - \frac{t_N^2}{2} \tag{15}$$

where G_N is tax revenue used as public spending in jurisdiction N (Baldwin and Krugman (2004)). This specification of the social welfare function is compatible with a government modeled as benevolent or as a Leviathan. If the government is benevolent, the utility function of the consumer should include public spending as argument; this is not the case here but other related papers have considered this possibility (López-Rodríguez et al. (2022)). The government budget constraint in N is:

$$G_N = \gamma t_N K \tag{16}$$

where t_N is the tax rates levied on capital by the jurisdiction N. Following Ottaviano and van Ypersele (2005), we set up without loss of generality that G is endogenously fixed and since the government budget must be balanced, the tax rate on capital is accordingly chosen. Consequently, the optimization problem is one-dimensional and the optimal tax rate on capital in N is obtained as a function of t_S, $t_N\left(t_S\right)$, that is, the reaction function of t_N with respect to t_S.

Coping with the symmetric optimization problem of the other jurisdiction is straightforward. The social welfare function is quite similar to that used for the jurisdiction

N, namely,

$$W_S = G_S - \frac{t_S^2}{2} \tag{17}$$

The government budget constraint is identical to the jurisdiction N (except from the obvious change in the sub-indexes). Therefore, using again the expression (14), the derivative of (17) with respect to t_S allow us to obtain its first-order condition. With the two reaction functions both local tax rates are achieved. For sake of brevity, given the symmetry of Nash-type game, we focus hereafter in the case of jurisdiction N.

Analytically, the derivative of t_N with respect to the agglomeration σ depends on constant parameters of utility function, the total endowment of capital K and the trade cost τ. That is, there exists a linear relationship between the local tax rate and agglomeration σ, which is present in the closed-form of such a derivative $(-\frac{2(b+cK)L\tau(-2a+b\tau)}{12b+6cK+c(b+cK)L\tau^2})$, being independent of σ. It can be also proved that this relationship can be zero or positive, given the restrictions on a, b, c and τ.

In order to explain our previous empirical results is critical to elucidate the sign of the above derivative. It is straightforward to see that the numerator turns out to be zero if $a = \frac{b\tau}{2}$. That is, given the trade cost τ, for relatively small intensity of preferences for the manufacturing good a in relation with the love-of-variety b, there is no impact of agglomeration on the local tax rates. But an increasing trade cost τ makes softer this condition. In other words, the higher the trade cost τ, the higher the relative preference for manufacturing good a might be for having a null slope in the relationship between agglomeration and local tax rates.

The intriguing issue here, not considered in the above model, is that heterogeneous consumer preferences across jurisdictions must be admitted for finding out different regimes in the connection between agglomeration and taxation. Indeed, for obtaining a positive slope in Figure 2 it is necessary to have consumer preferences for the manufacturing good even weaker than those in the jurisdictions with no impact of agglomeration on local tax rates, and null slope in Figure 2. Analytically, the condition becomes $a < \frac{b\tau}{2}$. Under such conditions, the agglomeration economies have positive effects on local taxation.

In this upward-slope stretch, the trade cost plays again a significant role. Given other things constant, the higher τ, the more likely having $a < \frac{b\tau}{2}$. Put differently, when the trade cost increases, the effect of agglomeration on the tax rate in the jurisdiction N goes up. We could even claim that the positive effect of agglomeration on local tax rates is reinforced as trade costs increase. The latter is a well-known result from the geographical economics literature (see for instance Baldwin et al. (2011)).

This is most easily seen for the cost-of-living effect. If the regions are very open in the sense that trade costs are low, then there will be very little difference in prices between the two regions whatever the spatial allocation of production is. Thus, shifting industrial production has only a minor impact on the relative cost of living. However, if trade is very costly, the share of varieties produced locally will have a big impact on price indices. Similar reasoning shows that the market potential advantage is more intense when trade costs are high, and the agglomeration economies become stronger by affecting local taxation.

A last suggestion to be taken into consideration in the explanation of empirical results is also related to the presence of heterogeneous preferences across jurisditions. Fréret and Maguain (2017) build a model on the standard features of horizontal tax competition theory initiated by Zodrow and Mieszkowski (1986), Wilson (1986) or Wildasin (1988) but adding agglomeration economies. This model is far away from our geographical economics framework but it achieves a similar result to ours, namely, two agglomeration regimes are found in a sample of French municipalities, with implications on the responses of local jurisdictions to changes in neighbouring business tax rates.

Their model includes heterogeneous consumer preferences across localities. Particularly, their preferences for the local publicly provided private good are more intense in the agglomerated area, assuming then a credible, positive correlation between the needs for public goods (for instance, infrastructure) and agglomeration forces. This is not our case but such a heterogeneity in preference could be indeed dealt with a local-specific social welfare function. Recall that we have assumed identical objective functions to be maximized by the local governments in the expressions (15) and (17). And new results, with thresholds in the relationship between agglomeration and local taxation, might well be found by departuring from this assimetry in the preferences for the public good, as in the Fréret and Maguain's (2017) contribution, but under the cordinates of geographical economics models.

7. Conclusions

This paper has paid attention on how the relationship between agglomeration and local tax rates works, focussing on whether thresholds effects are at play. With this aim in mind, we have studied the Spanish case, where a vast sample of municipalities over the period 2013-2020 has been exploited. Among the novelties included in the paper we highlight the use of advanced semi-parametric econometrics, not previously employed in this type of studies. Moreover, we have controlled for the potential presence of threshold effects in agglomeration economies on local business tax rates without imposing strong assumptions on the variables and their mechanics. For sure, other circumstances that usually are present in these analyses, such as endogeneity, spatial dependency and unobserved heterogeneity across jurisdictions, have been also taken into consideration.

Our approach has begun on the empirical ground to continue then with the fit of findings into standard theoretical models. The best model obtained after applying the semi-parametric techniques and a number of precautions supports the hypothesis of taxable agglomeration rents only for the larger municipalities. For the majority of jurisdictions, those whose market potential values are below the 75th-80th percentile, the relationship between agglomeration and local tax rates is found only when controlling for the endogeneity of spatial weights matrix, being significant at 10 percent level.

Additionally, for the larger jurisdictions in which the hypothesis of taxable agglomeration rents holds, the relationship is non-linear, an outcome that have not been previously found in the literature. In fact, we have proved that the standard linear assumption overestimates the impact of agglomeration on local tax rates decision-making, specially in municipalities with the highest values of market potential. In this context, the larger jurisdictions may take advantage of their ability to attract rents without lowering tax rates.

Both empirical results have been interpreted in light of theoretical models based on the standard assumptions of economic geography models (imperfect competition, increasing returns to scale and trade costs). In this context, we have reached a stylized characterization of the relationship between local tax setting and agglomeration. The simplest approach to such a link points to a linear relationship, as that found for the majority of municipalities. But, deepening on further connections among the variables, a chance for non-linearities arises.

Certainly, we achieve a condition ending in a positive and (potentially) increasing impact of agglomeration on local business tax rates. This condition involves consumer preferences (for manufacturing goods and love-of-variety) and trade costs. In general, increasing trade costs leads to positive impacts of agglomeration on the tax rates in municipalities.

Having arrived at this point, at least a couple of avenues for further research arise. One is precisely motivated by the theoretical fit of the above empirical findings we already did in the previous section. Although we have some understanding on the facts behind the presence of threshold effects between the local tax rates and the agglomeration, other factors deserve a further reconsideration. Particularly, we are referring to those involved in political economy extensions such as those coping with median voters or governments interested in maximizing the voters they receive in local elections.

Alternatively, we are also wondering on how our results might be modified if the jurisdictions play a different game. Instead of the Nash scheme we have followed here, what would it happen if one of the local governments behaves as a Stackelberg leader when sets its tax rates. Given a determined level of agglomeration, what the optimal behaviour of the other (follower) jurisdiction would be?

On the empirical ground, a straightforward extension of this paper includes considering alternative agglomeration measures to market potential (Balassa indexes, employment density). Other potentially fruitful avenues for further inquiry could study the effects of localization economies as opposed to the urbanization ones (the ones dealt with in this paper) based on the fact that in some industries firms tend to locate in a few areas creating highly specialized economic clusters.

In terms of policy implications some lessons can be drawn from the analysis carried out. Consumers preferences and trade cost (i.e., infrastructure policy) appear as key issues in local tax decision-making, beyond the standard approach of looking what the neighbors are doing. Additionally, as the hipothesis of taxable agglomeration rents seems to make sense once a determined level of potential market is achieved, the local policy-makers should be sure on which position their jurisdiction is in terms of agglomeration before taking decision on local tax business.

Acknowledgements

This research was supported by an FPU grant (Formación de Profesorado Universitario, FPU21/02763) from the Spanish Ministry of Universities (MIU) to Brais Pociña-Sanchez. Jesús López-Rodríguez and Diego Martínez-López also acknowledge the financial support from the Spanish Institute of Fiscal Studies.

We acknowledge the comments by Alberto Porto, Jorge Puig, Javier Suarez-Pandiello, participants in the XI Meetings of Local Public Finance in Latin America (Montevideo, 6-7 September 2022) and in the XXX Meetings on Public Economics (Badajoz, 4-5 May 2023). All remaining errors are solely our own responsibility.

References

Andersson, F. and Forslid, R. (2003). Tax competition and economic geography. *Journal of Public Economic Theory*, 5(2):279–303.

Baldwin, R., Forslid, R., Martin, P., Ottaviano, G., and Robert-Nicoud, F. (2011). *Economic geography and public policy*. Princeton University Press.

Baldwin, R. and Krugman, P. (2004). Agglomeration, integration and tax harmonization. *European Economic Review*, 48:1–23.

Basile, R., Durban, M., Minguez, R., Maria Montero, J., and Mur, J. (2014). Modeling regional economic dynamics: Spatial dependence, spatial heterogeneity and nonlinearities. *Journal of Economic Dynamics and Control*, 48:229–245.

Brülhart, M., Jametti, M., and Schmidheiny, K. (2012). Do agglomeration economies reduce the sensitivity of firm location to tax differentials? *The Economic Journal*, 122(563):1069–1093.

Brülhart, M., Bucovetsky, S., and Schmidheiny, K. (2015). Taxes in cities. In *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics*, pages 1123–1196. Elsevier.

Burdekin, R. C., Denzau, A. T., Keil, M. W., Sitthiyot, T., and Willett, T. D. (2004). When does inflation hurt economic growth? different nonlinearities for different economies. *Journal of Macroeconomics*, 26(3):519–532.

Charlot, S. and Paty, S. (2007). Market access effect and local tax setting: evidence from french panel data. *Journal of Economic Geography*, 7(3):247–263.

Charlot, S. and Paty, S. (2010). Do agglomeration forces strengthen tax interactions? *Urban Studies*, 47(5):1099–1116.

Christie, T. (2012). The effect of government spending on economic growth: testing the non-linear hypothesis. *Bulletin of Economic Research*, 66(2):183–204.

De Boor, C. (1978). *A practical guide to splines*. Applied mathematical sciences. Springer.

Devereux, M. P., Griffith, R., and Simpson, H. (2007). Firm location decisions, regional grants and agglomeration externalities. *Journal of Public Economics*, 91(3–4):413–435.

Eilers, P. and Marx, B. (1996). Flexible smoothing with b-splines and penalties. *Statistical Science*, 11(2):89–121.

Eilers, P., Marx, B., and Durban, M. (2015). Twenty years of p-splines. *SORT. Statistics and Operations Research Transactions*, (39):149–186.

Fréret, S. and Maguain, D. (2017). The effects of agglomeration on tax competition: evidence from a two-regime spatial panel model on French data. *International Tax and Public Finance*, 24(6):1100–1140.

Harris, C. D. (1954). The, market as a factor in the localization of industry in the united states. *Annals of the Association of American Geographers*, 44(4):315–348.

Jofre-Monseny, J. (2013). Is agglomeration taxable? *Journal of Economic Geography*, 13(1):177–201.

Jofre-Monseny, J. and Solé-Ollé, A. (2010). Tax differentials in intraregional firm location: Evidence from new manufacturing establishments in spanish municipalities. *Regional Studies*, 44(6):663–677.

Koh, H.-J., Riedel, N., and Böhm, T. (2013). Do governments tax agglomeration rents? *Journal of Urban Economics*, 75:92–106.

Lee, D.-J. and Durban, M. (2011). P-spline anova-type interaction models for spatio-temporal smoothing. *Statistical Modelling*, 11(1):49–69.

Luthi, E. and Schmidheiny, K. (2013). The effect of agglomeration size on local taxes. *Journal of Economic Geography*, 14(2):265–287.

López-Rodríguez, J., Pociña-Sánchez, B., and Varela-Candamio, L. (2022). Taxable agglomeration rents across spanish local labour markets. *mimeo*.

Mínguez, R., Basile, R., and Durbán, M. (2020). An alternative semiparametric model for spatial panel data. *Stat. Methods Appt.*, 29(4):669–708.

Minguez, R., Basile, R., and Durban, M. (2020). An alternative semiparametric model for spatial panel data. *Statistical Methods & Applications*, 29(4):669–708.

Minguez, R., Basile, R., Durban, M., and Espana-Heredia, G. (2022). Spatial and spatio-temporal semiparametric regression models with spatial lags (pspatreg).

Montero, J., Minguez, R., and Durban, M. (2012). Sar models with nonparametric spatial trends. a p-spline approach. *Estadística Española*, 54(177):89–111.

Muggeo, V. M. R. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine*, 22(19):3055–3071.

O'Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems. *Statistical Science*, 1(4):502–518.

Ottaviano, G. (2001). Home market effects and the (in) efficiency of international specialization. *mimeo*.

Ottaviano, G. I. and van Ypersele, T. (2005). Market size and tax competition. *Journal of International Economics*, 67(1):25–46.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics*, 11(4):735–757.

Sanso-Navarro, M. and Vera-Cabello, M. (2015). Non-linearities in regional growth: A non-parametric approach. *Papers in Regional Science*, 94:S19–S39.

Spanish Institute of Fiscal Studies (2018). Reform of the territorial financing systems. reports by the committees of experts of 2017. Technical report, Spanish Institute of Fiscal Studies.

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

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

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

Appendix

Figure 4: Non parametric detection of the threshold. Control estimation I (Column 3 of Table 2)

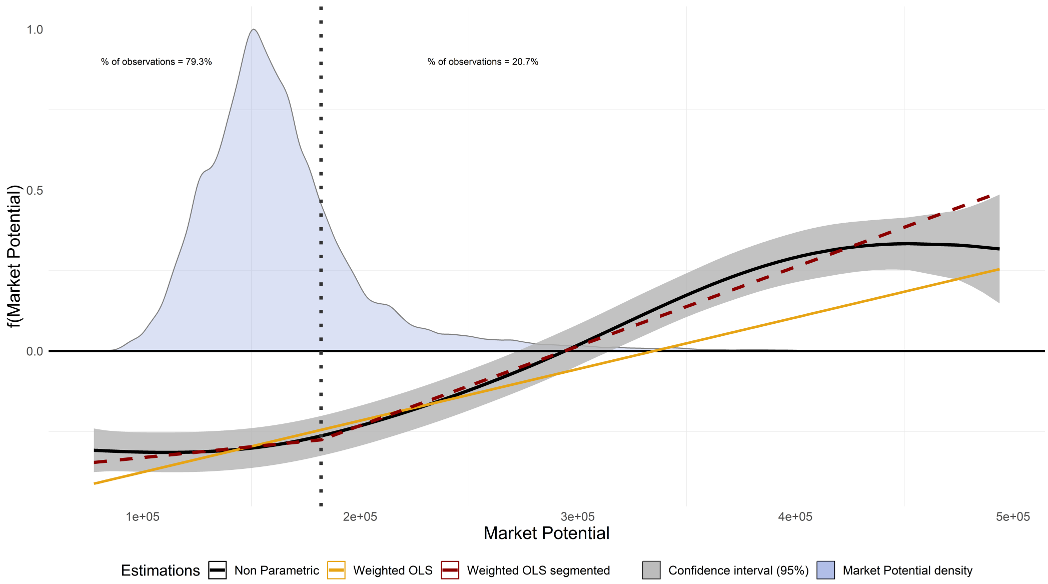


Figure 5: Non parametric detection of the threshold. Control estimation II (Column 4 of Table 2)

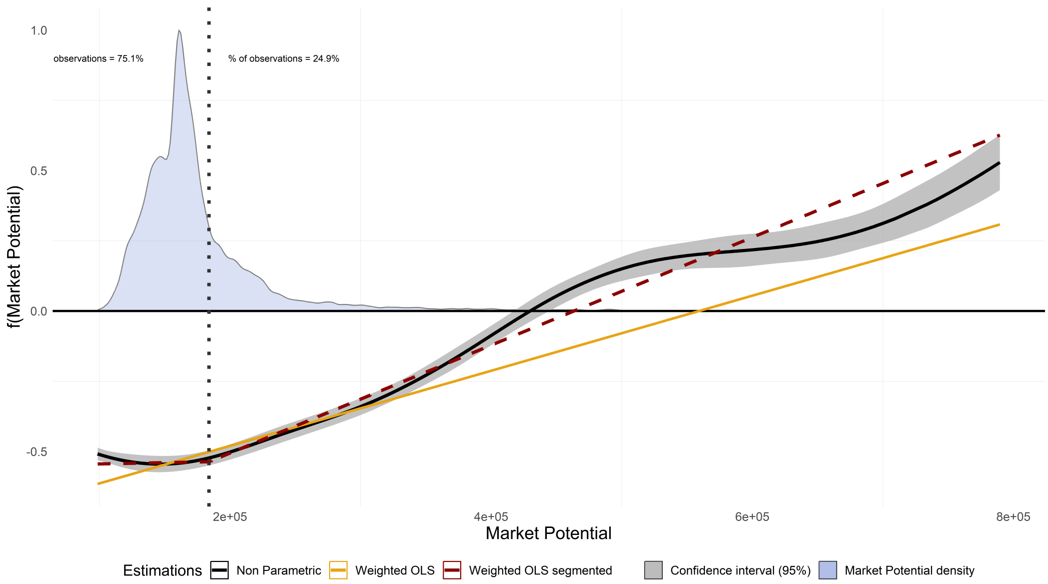


Figure 6: Non parametric detection of the threshold. Endogeneity of spatial weighting matrix (Column 5 of Table 2)

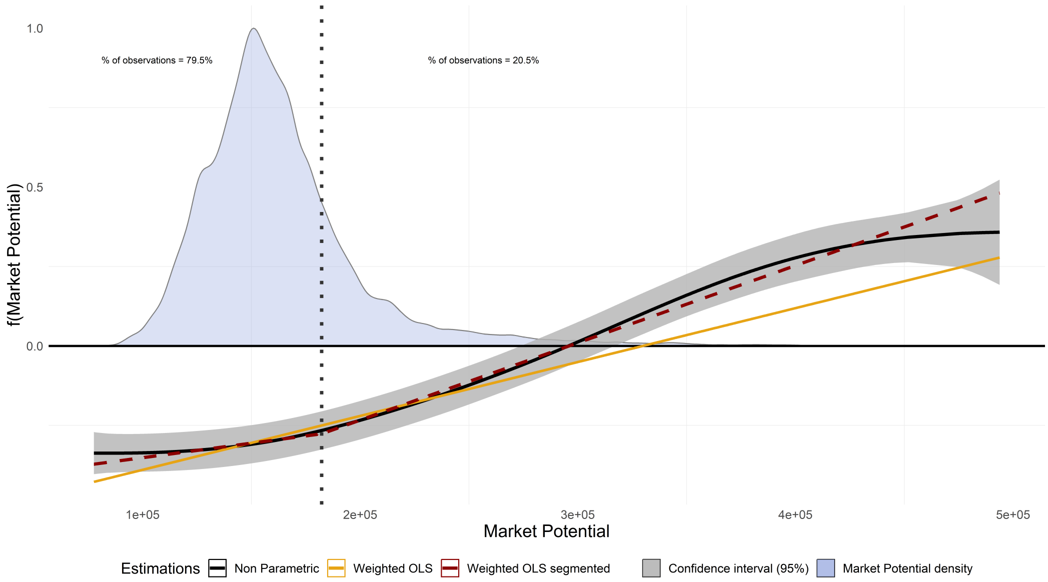


Figure 7: Non parametric detection of the threshold. Endogeneity if Market Potential (Column 6 of Table 2)