1 Introduction

Local government fragmentation raises several issues such as efficient provision of local public services. A common response to inefficient delivery is the implementation among municipalities of consolidation programmes. There are several examples of amalgamation in Europe especially among northern European countries. For example, in the case of Finland, the number of municipalities has decreased sharply (Saarimaa and Tukiainen, 2014) and in Denmark the number of municipalities was reduced from 271 to 98 in 2007. In recent decades, merger plans have been introduced by many countries around the world such as China (Hinnerich, 2009; Weese, 2015), Canada where the number of municipalities halved between 1996 and 2001 (Poel, 2000), and Israel where amalgamation reform was conducted in 2003 (Reingewertz, 2012).

An alternative to merging is creation of more flexible governance structures at the local level, that allow for joint provision of local public goods while maintaining jurisdictional autonomy (de Mello and Lago-Penas, 2013). Many countries (France, Spain, Italy, England, Ireland, The Netherlands, Portugal, Belgium, Brazil...) allow neighbouring municipalities to set up such inter-municipal agreement on voluntary basis, to provide local public goods to the residents of member jurisdictions (Dexia, 2012). In this setting, after joining an inter-municipal body within a functional cooperation, the municipalities continue to exist as entities.

In the theoretical literature, work on political economy includes several papers that focus on the optimal size of a coalition and its characteristics (see e.g. Bolton and Roland, 1997; Alesina and Spolaore, 1997, 2003, 2005, 2006; Brasington, 1999, 2003a, 2003b; Blume and Blume, 2007; Gordon and Knight, 2009; Hyytinen et al., 2014; Saarimaa and Tukiainen, 2014; Weese, 2015, Lago-Penas and Martinez-Vazquez, 2013). These studies show that nations are more likely to integrate if levels of income are comparable and voters have similar political preferences. In a within states context, there are similar issues, which are addressed mostly by work on fiscal federalism. [1] Most of the literature focuses mainly on municipality mergers, which following a coalition, cause municipalities to disappear, few papers investigate

1. Since the seminal article by Tiebout (1956), numerous articles show that fragmentation can have positive effects since small governments achieve a better matching between expenditure allocation and local preferences. Therefore, the problem of optimal size in the production of local public goods can modelled as arbitrage between the welfare gains expected from small localities and the economies of scale expected from the production of public services within a larger jurisdiction (Oates, 1972).

voluntary municipal cooperation (Hulst et al., 2009; de Mello and Lago-Penas, 2013; Di Porto et al., 2013).

To address this issue, we use the French cooperation experiment among local governments. We argue that if a municipality has a choice between remaining independent or joining a community, the political, economic, and socio-demographic environment may influence the cooperation decision in a variety of ways. Also, the cooperation decision is strongly dependent on what neighbours decide since by law, an inter-municipal community must include just contiguous localities. If no neighbours already cooperate, the probability of joining an inter-municipal community will be lower than if close neighbours already cooperating. This allows us to use a discrete choice model in which the decision to cooperate is affected by the spatial weighted average decisions of the municipality's neighbours. The main aim of the paper therefore, is to estimate the impact of neighbours' cooperation decisions on municipal decision-making, using a *spatial binary discrete choice model* and controlling for the political and socio-economic environment. [2]

The applied literature has rarely discussed cooperation decision-making. This is mostly due to the spatial econometric framework that the researcher should face, the complex computational issues related to estimation in this environment and the lack of reliable data on this phenomenon. This paper tries to deal with all these issues providing a robust estimation of a non linear spatial interaction model that is computationally simple and easily replicable.

Recent work on spatial interaction models discusses the validity and accuracy of traditional spatial econometric techniques (Gibbons and Overman, 2012) and provides new evidence on the advantages of instrumental variables (IV) methods over maximum likelihood (ML) technique (Lyytikainen, 2012; Baskaran, 2014). According to Gibbons and Overman (2012), weak identification in a spatial framework is often due to the use of non relevant instruments. [3] Exploiting exogenous variations from policy changes seems to be the most appropriate method to overcome these problems and derive valid exclusion restrictions. In a tax interaction model, Lyytikainen (2012) shows how the use of IV methods associated with careful choice of instruments can result in more reliable estimates of the spatial lag term in the Finnish context. Our paper follows Lyytikainen (2012) but advances the binary choice model case.

In the discrete choice model literature, the counterpart of the IV method is control function. Papke (2005) and Papke and Wooldridge (2008) have extensively investigated the panel data method for a fractional response dependent variable. In this analysis, the use of control function rather than other estimation methods has many advantages. Firstly, discrete choice models with spatial interaction are computationally intensive (LeSage, 2000; Pinkse and Slade, 1998). Therefore, maximum likelihood optimisation might be complicated because it requires the inversion of large matrices. When the number of

2. Note that Di Porto et al. (2013) estimate the relationship between the cooperation decision and the fiscal revenues raised to provide local public goods using a panel dataset of 35,000 French municipalities for 1995-2002. Their paper focuses on the role of the economic incentives to cooperate. However, they do not address the role of neighbouring municipalities decisions, nor the impact of political environment on the cooperation decisions.

3. The typical IV strategy uses spatial weighted covariates or spatial weighted lagged dependent variable as exclusion restriction.

observations grows, convergence can become unfeasible (Klier and McMillen, 2008). As an alternative, control function is a simple and well-known procedure, which is suited to large panel data, given that it involves only linear regressions and standard probit or Generalized Linear Model (GLM) methods in a binomial family. Also, control function can control for time invariant unobserved heterogeneity via Mundlak-Chamberlain devices (Papke and Wooldridge, 2008).

Following Papke and Wooldridge (2008), we use a control function to estimate a binary choice model with spatial interaction. In particular, we treat the spatial interaction as endogenously determined, and we solve it by imposing exclusion restrictions derived from the French institutional setting. Since finding a relevant instrument is crucial in a spatial setting, we provide a careful discussion of the weakness of our instrumental variable. Our analysis shows that our instrument based on the exogenous change in the French municipal block grant affects our endogenous variable in a way that is clearly different from linear. We find evidence that specifying our control function in a semi-parametric way improves the causal interpretation of our parameter and, therefore, provides a better identification strategy than the usual linear one. To our knowledge, this study is the first to use a control function in this context.

By applying our estimation strategy to panel data on French municipalities, we find that a municipality's decision to cooperate over the provision of local public goods depends on the decisions made by its neighbours. In particular, the probability of cooperating with neighbouring municipalities is higher if the latter already provide joint local public goods. We also investigate the role of political similarity, i.e. whether a similar political colour increases the propensity to cooperate of neighbouring jurisdictions. We find that a common political "colour" increases the incentive to cooperate but in a very weak way. This suggests that functional cooperation is likely to emerge due to a mimicking motivation and is poorly motivated by political alignment.

While achieving scale economies and internalizing spillover effects are clearly motivations for most municipalities to cooperate, own economic and demographic characteristics, such as a high level of fiscal revenues or a high proportion of elderly people, may lead municipalities to prefer isolation. Clearly, there is a trade-off between remaining isolated - which implies that public services are more tailored to citizens/voters - and the achievement of scale economies or the internalization of spillover effects within a large community.

We believe that this paper will contribute to ongoing debate on the reorganization of sub-national jurisdictions since it promotes the idea that the emergence of voluntary cooperation among local governments is less likely if municipalities have different socio-economic characteristics. In terms of public policies, the state cannot implement such agreements without allocating extra state grants to improve the incentives to cooperate for the richest local governments.

This paper is organised as follows. Section 2 discusses the conditions for voluntary cooperation among local governments. Section 3 presents the French institutional context. The econometric approach is discussed in section 4. Section 5 presents the results of our estimation and section 6 concludes.

2 Conditions for voluntary cooperation among local governments

A number of political economy models address the incentives to cooperate. [4] For example, Bolton and Roland (1997) and Alesina and Spolaore (1997, 2005, 2006) study the compromise in economies of scale that encourages cooperation and the problem of heterogeneity of preferences that promotes dissolution, which together determines the optimal size of a given coalition. Nations are shown to be more likely to integrate if they have comparable levels of income and their voters share close political preferences.

Within a state, Oates (1972) points out that the formation of a coalition internalizes inter-jurisdictional spillovers externalities and can be a solution to the problem of free riding. The motivation for coalition formation within a country has also attracted considerable interest from theoretical and empirical economists. Borge and Rattso (2002) confirm that in Norway, political strength, measured by the fragmentation of parties and types of coalition, is important for growth and influences decentralized spending and the formation of agreements. Also for Norway, Sorensen (2006) shows the impact of political transaction costs on voluntary local government mergers. He identifies the role of state grants, revenue disparities and expected changes to party strengths after unification, in merging process. Gordon and Knight (2009) and Weese (2015) propose structural econometric methods to analyse spatial mergers. Where a merger is decided voluntarily by the municipal councils, they show that the local political environment plays an important role in the mergers decision. These authors point out also that political fragmentation can be an obstacle to the formation of coalitions and conclude that the political environment can discourage optimal coalition formation. Saarimaa and Tukiainen (2014), using Finnish municipal data, study the case of mergers that were decided independently at the local level, and analyse empirically the coalition formation of local governments. Using a novel reduced form econometric procedure that allows for multi-partner mergers, they find that the decision to cooperate or merge depends on the preferences of voters, and that mergers are much less likely if the distance between median voter in the coalition and the center is large. They argue that in order to reduce political competition and to be re-elected, councillors seems to prefer mergers. Finally, Hyytinen et al. (2014) study how anticipated changes in the electoral vulnerability of municipal councillors also affect their voting behaviour over municipal mergers.

Since we want to address a specific class of cooperation where municipalities continue to exist as entities

4. Issues related to coalition formation and analysis of bargaining are the subject of number of theoretical studies of game theory, which are beyond the scope of the present study (see e.g. Myerson, 1977; Shenoy, 1979; Aumann and Myerson, 1988; Bloch et al., 2006).

after joining an inter-municipal body, we address three factors that are key to functional cooperation decision-making: economic, political and socio-demographic determinants.

Although improved cost efficiency is a major incentive for consolidation, it is one of the most difficult key elements for the municipality to assess a priori (Sorensen, 2006). To assess the trade-off between remaining isolated and cooperating, a municipality will compare its revenue with what can be expected after joining the inter-municipal community. Therefore, we expect that if a municipality has higher levels of (exogenous) revenues than her potential partners, it may be opposed to cooperation to avoid revenue sharing with poorer municipalities. Inversely, a relatively poor (in terms of fiscal revenues) municipality will have more incentives to share the resources with richer neighbours.

Citizens of neighbouring municipalities may have more or less close policy preferences. The political costs associated with larger, more heterogeneous jurisdictions have been the subject of long debate (Sorensen, 2006). The new local official in the case of merger may pursue a different policy to those pursued by former separates localities, leading to loss of allocation effectiveness (Oates, 1972). In the case of functional cooperation where former jurisdictions do not disappear, the political factor may be less strategic than in the case of merger case. Although cooperation among local governments to provide joint local public goods is also more likely if voters share close political preferences, each municipality retains autonomy for the provision of specific municipal public goods. A cooperating municipality can transfer some spending to the inter-municipal community but retain many important competences (such as local urban services, buildings, nursery provision, primary schools, and municipal roads maintenance), which respond to specific local preferences.

The socio-demographic characteristics of municipalities are closely related to economic and political determinants. First, the larger the municipality's population, the greater will be its local public needs and the greater will be the spillovers of public goods to neighbouring localities. Free riding behaviour from citizens who do not live in a city but benefit from its public good delivery provides a dense jurisdiction with a greater incentive to share or transfer to the inter-municipal group some competences and associated supply of services. Second, the prospect of shared social responsibilities could also be a strong motive for some municipalities to join an inter-municipal agreement. However, the provision of social services is not ordained by law for inter-municipal communities. Moreover, elderly people are less sensitive to investment by inter-municipal jurisdictions (sport facilities, public transport, road investment). A high proportion of elderly people in one municipality is likely to influence the probability of cooperation negatively.

Finally, we can conclude that if a municipality has a choice between remaining independent or joining a community, the political, economic, and socio-demographic environment may influence the cooperation decision in a variety of ways. Neighbourhood characteristics in terms of economics (e.g. per capita fiscal revenues), politics (political affiliation) and demography are key factors that need to be addressed to understand the cooperation decision. In this context, the French experience is an interesting case.

3 The French institutional context

France is a very fragmented country; since the lowest layer of local government consists of more than 36,700 municipalities and the French local institutional context is characterized by three overlapping tiers of local government. The middle tier consists of 96 counties or *départements*; the top tier consists of the 22 French regions. Municipalities are responsible for local urban services, buildings, provision of nurseries, primary schools, sports facilities, maintenance of municipal roads, and urban public transport. The counties administer social assistance; and maintain county roads and middle schools. Regions are responsible for provision of vocational training, economic development, and building and maintenance of high schools.

Local revenues come mainly from taxation (54%), block grants (23%), and borrowing. The local business tax (*Taxe Professionnelle*) [5] is the major source of local government tax revenues, accounting for approximately 45% of the revenue derived from direct local taxes. The tax base consists mainly of capital goods and is related to the rental values of buildings, and of equipment (assumed to be 16% of the equipment cost). The remaining tax revenues are collected from households in the form of residential tax (*taxe d'habitation*), property tax (*taxe foncière sur le bâti*), and land tax (*taxe foncière sur le non bâti"*). All municipalities receive a state grant, related to various criteria including population and tax bases.

Since the beginning of the 1990s, several laws on local cooperation have been passed to solve the problem of municipal fragmentation:

— The law of February 6, 1992 established the basis of inter-municipal cooperation in order to promote economic development.

— The law of July 12, 1999 known as the "*Chevнement law*", is one of the main laws encouraging consolidation of inter-municipal cooperation in France.

— A government reform, in 2010, enforced all municipalities to join an inter-municipal structure before January 1, 2014.

Local cooperation is expected to *i)* reduce local public spending through the achievement of substantial economies of scale in the production of some local public goods such as public transport, cultural and sports facilities etc. (mainly investment spending) and *ii)* limit fiscal and spending inequalities among member municipalities. These objectives are achieved through the transfer of some tax and spending powers from the municipalities to the corresponding inter-municipal group.

5. This tax, which was related mainly to private capital, was removed in 2010 and replaced by a territorial economic contribution based on property and firm value added.

In France, responsibility for creating inter-municipal communities (*'Etablissement public de coopération intercommunale' or EPCI*) is left to the municipalities, which (before 2010) could cooperate with the municipality(ies) of their choice, provided that they share at least one common border. This is "the principle of territorial contiguity". The EPCI is governed by a board of delegates elected by municipal councils among their members. Therefore, unlike council members in municipalities, inter-municipal officials are not directly elected by population. [6] The number of seats held by a municipality is generally proportional to the municipal population. Each municipality must have at least one seat, and no single municipality can hold more than half of the inter-municipal council seats.

Cooperation among local governments has been widely promoted by government based on financial incentives. In 1999, the inter-municipal community was awarded a new state unconditional transfer (based mainly on community population and inter-municipal tax bases), which coexists with the block grant attributed to the municipality. The latter remains unchanged whether the locality cooperates or not and is based on a complex computation that relies on municipal population and tax base.

The communities also benefit from specific tax revenues since they apply additional tax rates to the four municipal taxes (business, residential, property and land taxes) or set a single business tax (SBT) or *Taxe Professionnelle Unique rate.* [7] If the municipality chooses to impose a SBT, it loses the right to set her own business tax rate but can still set the rate for the three household taxes (residential, property and land taxes).

We employ a panel data set of 1,056 municipalities in the French region - Brittany - over a period of 8 years (1995-2002). The choice of this region is justified by its geographical position: Brittany occupies the northwest peninsula of continental Europe in northwest France. It is bordered by the English Channel to the north, and the Atlantic Ocean to the west and south. Thus, Brittany is affected by its borders with other regions to only a small extent. Note also that there is no specificity in this region concerning inter-municipal communities since the laws on local cooperation are national laws.

Figures 1 and 2 depict the municipalities that entered into an inter-municipal agreement between 1996 and 2002. Inter-municipal cooperation in this period was quite successful. However, some municipalities remained independent. The obvious financial gains from cooperation (through the extra state transfer) and the provision of new public services might be counterbalanced by each cooperating municipality losing some fiscal autonomy through devolution of competences and tax revenues to the community. Our aim is to explain why some municipalities decided to remain independent while others concluded joint local public goods provision agreements.

6. Note that this is not the case anymore since 2015. Citizens vote directly for specific inter-municipal officials at municipal elections.

7. During the period of our study, there was a third - marginal - case, which was a mixed case where the community sets a single business tax rate and an additional tax rate on households. The community received tax revenue from households and business while municipality i sets a tax only on households.

Figure 1

Figure 2

4 Econometric approach

Since by law, French local governments were free to decide to join an inter-municipal group, provided that they share a common border, this allows us to use a discrete choice model. In addition, municipalities continue to exist as entities after joining an inter-municipal body (Di Porto et al., 2013). Therefore, the starting point of our econometric approach is a standard binary discrete choice model where the dependent variable C_{it} is given by:

$$C_{it} = \begin{cases} 1 & \text{if municipality } i \text{ cooperates at time } t \\ 0 & \text{otherwise} \end{cases}$$

The reduced form of our panel data model is:

$$C_{it} = \phi(\beta X_{it}^{'} + u_{it}) \tag{1}$$

where i is the index of municipalities, t is an index of the time dimension, X_{it} is the matrix of the covariates, β is the vector of coefficient, u_{it} is the error term, and ϕ is a normal distribution.

Allowing for the possibility that the error term includes unobserved fixed municipal characteristics, the model can be written as:

$$C_{it} = \phi(\beta X_{it}^{'} + \eta_i + e_{it}) \tag{2}$$

where $\eta_i = \bar{x}_i = T^{-1} \sum^T x_{it}$ are Mundlak-Chamberlain transformations controlling for the unobserved time-invariant individual effects and e_{it} is an independently and identically distributed error term. Papke and Wooldridge (2008) show that under conditional normality assumptions, the estimation of fixed effect using covariates time average is straightforward.

To assess the neighbourhood effects on the decision to cooperate, we consider a spatial specification for our panel probit model including the spatial lag of the dependent variable as a regressor, $\sum_{j=1}^{N} w_{ij} C_{jt}$, i.e. the weighted average on its neighbours' values. Neighbouring municipalities are defined in subsection 4.2.

The spatial probit model with fixed effects is therefore given by:

$$C_{it} = \phi(\rho \sum_{j=1}^{N} w_{ij} C_{jt} + \beta X_{it}^{'} + \eta_i + e_{it}); \tag{3}$$

where N is the number of municipalities, w_{ij} is an element of the weight matrix W indicating whether observations i and j are neighbours, and $\sum_{j=1}^{N} w_{ij} C_{jt}$ is the spatial lag of the dependent variable. We choose to estimate a spatial lag model but also a Durbin model specification. [8]

The form of the spatial Durbin model becomes:

$$C_{it} = \phi(\rho \sum_{j=1}^{N} w_{ij} C_{jt} + \beta X_{it}^{'} + \gamma \sum_{j=1}^{N} w_{ij} X_{jt}^{'} + \eta_i + e_{it}) \tag{4}$$

where $\sum_{j=1}^{N} w_{ij} X_{it}^{'}$ is the matrix of spatially lagged explanatory variables with associated parameter vector γ. In particular, the coefficient β measures the influence of municipality' i's characteristics on municipality i's decision to cooperate, while the coefficient γ measures the impact of neighbours' characteristics on i's cooperation decision. In both equations (3) and (4) we include in $X_{it}^{'}$ year dummies in order to control for year fixed effects.

We estimate equation (3) and (4) using different techniques and within different specifications. Firstly, we estimate a fixed effect spatial lag (Durbin) model using a linear probability model (LPM). In this case we remove η_i, which are redundant given that we control for municipalities fixed effects using demeaned variables (see Elhorst, 2003).

Secondly, we perform the linearised GMM (L-GMM) procedure in Klier and McMillen (2008). The technique is intriguing since the model in principle performs a non linear estimation and, is first linearised around the convenient starting point $\rho = 0$. In this way it is possible to perform a standard probit model followed by two-step least squares to obtain consistent estimates of the spatial lag coefficient. This procedure uses traditional $\sum_{j=1}^{N} w_{ij}^2 X_{jt}^{'}$ as exclusions restriction (where w_{ij}^2 are the elements of the W^2 matrix) and can be extended to the panel case including Mundlak-Chamberlain devices.

Our last specification is probably the simplest in terms of computation. We apply control function (CF) method considering $\sum_{j=1}^{N} w_{ij} C_{jt}$ as an endogenously determined variable. In particular, we use a two-step procedure where: i) firstly, we regress the endogenous variable $\sum_{j=1}^{N} w_{ij} C_{jt}$ on instruments, exogenous covariates and time average covariates, and we obtain the residuals $\hat{\nu}_{it}$; ii), we include those residuals in the original regression, that is (for the spatial lag case):

8. As observed by Gibbons and Overman (2012), Durbin and spatial error models are equivalent under a set of nonlinear common factor constraints on the coefficients.

$$C_{it} = \phi(\rho \sum_{j=1}^{N} w_{ij} C_{jt} + \beta X_{it}^{'} + \eta_i + e_{it} + \hat{\nu}_{it}) \quad (5)$$

This last model can be estimated as a Generalised Linear Model (binomial family) and under fairly standard conditions produces consistent estimate of ρ. Given that the second-step regression involves a generated regressor, i.e. the first stage residuals, we follow Papke and Wooldridge (2008) and compute the standard errors for the average partial effect via bootstrap. [9]

Before discussing our results, we explain the identification strategy in more detail (subsection 4.1), discuss neighbours decisions and other control variables in subsection 4.2 and finally present identification issues in subsection 4.3.

4.1 Identification strategy

According to Gibbons and Overman (2012), weak identification in a spatial framework is often due to the use of the traditional $\sum_{j=1}^{N} w_{ij}^2 X_{jt}^{'}$ as exclusion restrictions. Exploiting exogenous variations from policy changes, therefore seems the most appropriate method to derive valid exclusion restrictions in tax competition setting (Lyytikainen, 2012). This is still more advisable when the dependent variable is a discrete binary variable, where the little independent variation due to multi-collinearity may cause serious estimation problems.

The event of the census in France provide the possibility to develop a credible instrument for our endogenous variable based on municipal state block grant. Although there is an unconditional transfer allocated to the inter-municipal community, state block grant is a part of the municipal revenues always levied by the municipality regardless the decision to cooperate or not. The extent of this block grant - as a part of local tax revenues - is likely to influence the decision to remain isolated or not. Therefore, in our case, block grant itself is not a good candidate for a valid instrument. [10]

Differently, in order to achieve identification for our spatial interaction term, we follow Gordon (2004) and we use census discontinuity in block grant computation as exclusion restriction. State (block) grant is based on a complex computation, that relies on municipal population and tax base. These variables are therefore updated at every census (1990 and 1999). As explained in Gordon (2004), an exogenous discontinuity provided by the update may arise in the distribution of state grants. The unpredictability of both the extent and the direction of the updating in the key criteria (municipal population and tax base) for French municipalities is thus exploited to ensure exogeneity.

9. The computation of the bootstrap standard error in a spatial panel context is not well established in the literature yet. However, we provide a new methodology which is suitable in this case (see Appendix A for details).

10. The use of a state grant (block grant) allocated to municipalities as an exclusion restriction in empirical local public finance is extensive. Papke (2005) and Papke and Wooldridge (2008) provide examples of control function estimation with block grant as instruments.

In particular, and similarly to Gordon, we build the discontinuity in the following way. First, we regress state grants on lagged state grants, population, local business tax base (which is the main source of tax revenues for municipalities) and we include municipal and year fixed effects. This regression is performed for the period 1993 to 1998, in which the 1990 census data were used to compute state grant, and is then used to predict the values for 1999 to 2002. These predictions can be considered as a sort of synthetic control values or counterfactual values of block grants, derived from the 1990 census information. Secondly, we derive our census discontinuity in the block grant by computing the difference between actual state grant values and these predicted values. The discontinuity therefore is equal to 0 from 1993 to 1998 while from 1999 to 2002 is equal to the difference between the expected grant based on 1990 census and the grant actually received by each municipality. This procedure allows us to compute a non-expected change in block grants that is likely to influence the cooperation decision-making. Municipalities being richer than expected may have an incentive to remain isolated, whereas those who did not anticipate a loss in their block grant may be more likely to try to get new revenues through the state transfer, which is allocated to the inter municipal community.

Consequently, if BGC_{it} is our indicator of unexpected block grant change, $\sum_{j=1}^{N} w_{ij} Z_{it} = \sum_{j=1}^{N} w_{ij} BGC_{jt}$ will be our exclusion restriction for $\sum_{j=1}^{N} w_{ij} C_{jt}$. It should be noted that $\sum_{j=1}^{N} w_{ij} Z_{it}$, which is a weighted average indicator, is a suitable instrument for the spatial correlation coefficient. Intuitively this means that the spatial endogenous component is predicted by a spatial process that is exogenously determined by the census discontinuity event, exactly as in Lyytikainen (2012). De facto, some of the cooperating neighbours in each municipality have chosen cooperation for exogenous motivations, i.e. because their expected revenue was altered by an unpredictable event.

While exogeneity is ensured by the discontinuity at census year, relevance should be tested with the proper statistical tools. Figure 3 represents a non parametric interpolation between $\sum_{j=1}^{N} w_{ij} Z_{it}$ (horizontal axis) and $\sum_{j=1}^{N} w_{ij} C_{it}$ (vertical axis).

Figure 3

As shown in Figure 3, the relation between the weighted average cooperation and the weighted average discontinuity variable is highly non linear and, in particular, almost constant for high negative values of the unexpected block grant change (i.e., for municipalities who unexpectedly get a huge loss in their revenues after the census), increasing when the unexpected loss decreases and, finally, strongly negative for municipalities who unexpectedly get a gain in the block grant. The economic interpretation of such pattern is clear: a municipality that gains some block grant unexpectedly is less likely to cooperate than a municipality that unexpectedly loses some revenues after the census change. Yet, this relationship is clearly not linear. Since a first stage regression that expresses such relation in linear term may mis-specify the underlined model and may lead to a weak instrument, in the first stage regression of the control function we consider both a linear relationship and a semi-parametric one, in order not to lose precious information in the prediction of the residual term. In both cases we test for weak instruments to detect the possible mis-specification and improve the causal interpretation of the parameter. [11]

4.2 Neighboring decisions and other control variables

In order to estimate the impact of neighbours cooperation decisions on municipal decision-making, and following the literature on voluntary cooperation decisions, the cooperation decision (a discrete choice variable taking the value 1 in the event of voluntary cooperation) is regressed on its spatially lagged value (with spatial weights given by either contiguity or geo-political distance), and on a set of control variables proxying for economic, socio-demographic, and political characteristics. Descriptive statistics are reported in Table A1 in Appendix B.

First, the cooperation decision is strongly dependent on what neighbours decide since by law, an inter-municipal community must include just contiguous localities. If no neighbours cooperate, the probability of joining an inter-municipal community will be lower than if close neighbours already cooperating. We can also expect some mimicking behavior related to cooperation from local officials who might believe it would be stigmatising to remain isolated when most neighbouring localities have signed an intergovernmental agreement. Citizens/voters may become aware of new public services provided by a community (e.g. public transport, cultural and sports facilities) and may put pressure on their local government to join the community. [12] Officials might recognize that remaining isolated from an existing close community (which provides a range of good public services to firms and households) could lead to capital flight from the territory. Tax base mobility might explain the propensity of officials to imitate the cooperation decisions of neighbours. [13] It is likely also that municipalities, when deciding whether or not to cooperate, will mimic neighbours' behavior following the trend suggested by Manski (1993). We can expect also that obtaining information on cooperation (expected revenues, expected state grant, etc.) will be easier for an independent municipality located next to a group of cooperating municipalities. Mimicking behaviour related to cooperation is likely to be observed in this context. The literature on local fiscal decision-making provides extensive developments on this kind of spatial spillover and its identification (Brueckner, 2003).

As commonly used in this literature, the spatial weight matrix W used in our analysis is a row-standardised matrix based on contiguity between two municipalities i and j. The value of the element

11. The instrument is used in addition to time lagged covariates in the first stage regression. This ensures that the instrument is not correlated with some unobservable time-varying characteristics through the covariates.
12. The argument is similar to a yardstick competition models (Salmon, 1987; Besley and Case, 1995) where incumbents imitate the public decisions of their neighbours to stand for reelection.
13. This is a tax competition argument.

w_{ij} of W is given by:

$$w_{ij} = w^*_{ij} / \sum_j w^*_{ij}$$
$$w^*_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } i \text{ and } j \text{ share a common border} \\ 0 & \text{if } i \text{ and } j \text{ do not share a common border.} \end{cases} \quad (6)$$

We are aware also that proximity in term of distance is not the only factor in our case. Citizens of neighbouring localities do not necessarily have the same policy preferences. As argued before, where a municipality has a choice between remaining isolated or joining a community, politics may have an impact on the cooperation decision in a variety of ways. If a municipality is surrounded by independent localities, the creation of an inter-municipal community is more likely to happen if the neighbouring local officials are of the same political belief. For a municipality that is spatially contiguous to an existing community, a positive decision to join this community will be more likely if the president of the inter-municipal council (elected by local councillors from member municipalities) belongs to the municipality's ruling party. However, in France, functional cooperation is different from a merger: the cooperating municipalities continue to exist as entities after joining an inter-municipal body and the cooperating municipality mayors remain in post after the creation of the inter-municipal jurisdiction. They retain some bargaining power within the community and some strategic competences within their particular municipal territory (such as local urban services, buildings, nurseries, primary schools, sports facilities). Therefore, political factors may be less important in the French context compared to the northern European countries.

Since more than 80% of French municipalities are very small (less than 2,000 inhabitants), many French mayors do not have a political affiliation and do not want to be associated with a particular political party. Therefore, the political colour of the municipal government is not directly observable and there is no political data set for these small municipalities. To proxy for political affiliation, we use the results of the first round of the presidential elections since these are more likely to represent real affiliation. This variable comes from the board of elections ("Bureau des Élections et Études politiques, Ministere de l'Interieur"). During our period of study, there were two municipal elections (in 1995 and in 2002) but only one presidential election in 1995. We therefore consider a constant municipal political affiliation over time. In 1995, 546 municipalities could be considered right-wing and 510 localities left-wing.

To capture the political environment of each municipality i, we construct a geo-political matrix as follows.

1. For each municipality i we construct a political colour variable PC_i as:

$$PC_i = \begin{cases} 0.5 & \text{if municipality } i \text{ voted for a right-wing candidate} \\ -0.5 & \text{if municipality } i \text{ voted for a left-wing candidate.} \end{cases}$$

2. Then we built a political matrix P_{ij} such that:

$$P_{ij} = \begin{cases} 1 & \text{if } PC_i = PC_j \\ 0 & \text{otherwise.} \end{cases}$$

3. Finally, we interact the political matrix P with the contiguity matrix W. In particular, for any couple of municipalities i and j, the value of the element w_{ij}^{geoPol} of the geo-political matrix is given by:

$$\begin{aligned} w_{ij}^{geoPol} &= w_{ij}^{geoPol^*} / \sum_j w_{ij}^{geoPol^*} \\ w_{ij}^{geoPol^*} &= \begin{cases} 1 & \text{if } i \neq j,\ PC_i = PC_j \text{ and } i \text{ and } j \text{ share a common border} \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{7}$$

To assess the trade-off between remaining isolated and cooperating, a municipality will compare its revenue with what can be expected after joining the inter-municipal community (Sorensen, 2006). However, efficiency improvements in the production costs of local public goods may be difficult to evaluate. Therefore, if a locality has a higher level of fiscal revenues (per capita) than its potential partners, it will be less likely to cooperate. We argue that a municipality makes its decision at time t on the basis of the level of its fiscal revenues (i.e. the sum of its tax revenues and the per capita state grant) at time $t-1$. We expect that a high level of fiscal revenues at time $t-1$ will reduce the probability of joining the community whereas a financially poor locality will have a greater incentives to forge an agreement with neighbouring localities. The obvious financial gains from cooperation (from the extra state grant) and the provision of extra public goods are counterbalanced by each cooperating municipality's loss of some fiscal autonomy through devolution of competences and tax revenues to the community. In entering a cooperative agreement with other localities, a relatively "rich" municipality loses control over a part of its revenues, which are redistributed among the entire community.

We also include the following explanatory variables: population density, and share of elderly people in municipality i. We expect a positive impact of density for the following reason. First, the denser the municipality, the greater will be the local public needs and the public goods spillovers to neighbouring localities. In order to reduce free-riding behaviour from citizens who do not live in a city but benefit from her public good provision, the municipality will have an incentive to share or to transfer some competences and the associated supply of services, to the inter-municipal group. We would expect a negative sign for the share of elderly people since this category of the population has comparatively lower needs related to this investment spending, which are financed by the inter-municipal community. However, if the existence of a community is seen as a solution to sharing or transferring the supply of other specific services to this population category, we should observe a positive impact. Since social services are not legally compulsory for inter-municipal communities, the expected sign of this variable is ambiguous. [14]

As usual in this spatial economic context, we cannot rule out the possibility that voluntary cooperation might be strongly dependent on the weighted average of neighbours' characteristics. We perform a Durbin model specification in which the above socio-economic characteristics of neighbours are spatial lagged covariates. We expect that being surrounded by richer municipalities (e.g. with higher per capita fiscal revenues) would reduce the probability of cooperation since rich localities are not likely to transfer their resources to the community. The lagged covariate associated with elderly might exhibit an ambiguous sign for the reasons given above. Finally, a dense neighbourhood might be a strong incentive to sign an inter-municipal agreement since local government cooperation allows internalization of spillover effects in the supply of public goods.

4.3 Identification Issues

As Gordon and Knight (2009) and Weese (2015) argue, the researcher faces a number of issues when investigating cooperation among local governments: multi-sided decisions, the need for a multinomial approach and the possible violation of the stable unit of treatment values assumption (SUTVA) through spatial dependency. However, in our context, we believe that our econometric approach deals with these issues as detailed below.

First, as mentioned by the referee, the issue of multi-sided decisions does not apply in our context since the cooperation decision is one-sided and has to be accepted by partners if there is territorial contiguity. This is an important difference between functional cooperation and merger. Therefore, the functional cooperation framework helps our estimation since the municipality remains in charge and, therefore, can be used as an observation in a balanced panel.

The second issue is related to the set of potential partners, which would require a multinomial approach. However, we can consider the cooperation decision-making process represented by a decision tree where the municipality first decides to cooperate (or not) and second chooses its partner if cooperation was chosen in the first stage. Our paper studies only the first stage of this decision process, i.e. to join/not

14. Due to collinearity problems, we had to remove from the estimations population, percentage of unemployed and share of young people in the municipality i.

join a community. A multinomial approach would be necessary in order to study the second stage of the decision process. Note that we find that a jurisdiction faces a high (close to 7) average number of alternatives in our context. We are aware that our paper is less informative than a multinomial approach which would investigate the partner choice, but we believe that our strategy remains relevant for the first stage decision. We provide information on the decision to join (or not) a functional cooperation and on the determinants of this choice.

The last issue deals with spatial dependency and with the possibility that in our framework the so called stable unit of treatment value assumptions (SUTVA) are not verified. Our instrumental variable strategy can be seen as an attempt to develop a quasi experiment since we try to estimate a local average treatment effect (LATE). The instrument $\sum_{j=1}^{N} w_{ij} Z_{it}$ can be effectively seen as a variable that randomizes the treatment $\sum_{j=1}^{N} w_{ij} C_{it}$. As a consequence, this exogenous change affects C_{it} through $\sum_{j=1}^{N} w_{ij} C_{it}$ and allows us to identify the spatial lag coefficient which, in turn, would estimate the LATE. In this case, the SUTVA should guarantee that the effect of a change in $\sum_{j=1}^{N} w_{ij} Z_{it}$ affects the specific local government decision of a locality i through the changes in its $\sum_{j=1}^{N} w_{ij} C_{it}$ specific value. If this is not the case the SUTVA is violated. [15]

However, SUTVA violation is hard to identify in most social interaction models. The fact that we only address the cooperation/non cooperation decision rather than a multi-sided decision of partner helps us to achieve identification since the choice set is simpler.

As Mansky (2013) argues, the SUTVA is quite restrictive in situations that allow for social interactions. Our baseline specification allows a potential treatment to depend on both the individual treatment and the average level of treatment in the group of neighbours (as captured by $\sum_{j=1}^{N} w_{ij} C_{it}$ and also by $\sum_{j=1}^{N} w_{ij} X_{it}$ in the Durbin model). Mansky defines this as a functional interaction response, a relaxed version of the SUTVA or "individualistic treatment response". In other words, the social interaction can happen through some function of the distribution of the treatments across the hypothetical specific groups. In our case, there are no well-defined specific groups, but intuitively these are represented by the spatially weighted average covariates. Relaxing SUTVA in our case would lead to what Mansky calls distributional interaction, where individual treatment response depends on the distribution of treatments across others in the group, but not on the size of the group or the identity of those treated (Gibbons et al., 2015). Other possible relaxations of the SUTVA are plausible, and progressively weaker assumptions make the identification less credible. In our case, the lower levels of the estimated coefficients using a more

15. An example of violation would be as follows: suppose that a change in $\sum_{j=1}^{N} w_{ij} Z_{it}$ of a non contiguous municipality (i.e. not contiguous with i) changes the cooperation status of one neighbour of the local government i, the cooperation decision of the municipality i could be affected by the cooperation decision of this neighbour and not directly by the exogenous change. In other words, a spillover induced by a treatment of a municipality, which is not contiguous to i, changes the treatment status of local government i and then affects its outcome. If this violation happens, we can no longer be certain that the estimated LATE coefficient is well identified.

credible instrumental strategy let us believe that the parameter is relatively better identified respect to the traditional works in this field. Note that Gibbons et al. (2015) recognize that the literature on this issue is in its infancy. Mansky (2013) also states that outside of linear models, the relationship between identification of structural functions and reduced forms remains largely an open question.

5 Results

Tables 1 and 2 report respectively the estimations for the Spatial Lag Model in Eq. (3) and the Spatial Durbin Model in Eq. (4), using the linear probability model (LPM), the L-GMM probit and the control function (CF) approach, with the contiguity matrix and the geo-political matrix. The L-GMM and CF estimators are performed using either W^2X or our external instrument WZ. Results using WZ in a semi-parametric way are exhibited in each Table after $s(WZ)$. Note that ρ is the spatial lag coefficient while $\hat{\nu}$ is the predicted residual computed after the first stage regression using the control function approach. For the sake of space, first stage regressions are shown in Tables A2 and A3 in Appendix C. Full models estimations are reported in Tables A4 to A7. Table A5 shows a linear IV specification (using a contiguity matrix only).

Table 1: Spatial lag estimation results

Table 2: Spatial Durbin estimation results

The estimation results in Tables 1 and 2 show that the spatial dependence parameter (ρ) is almost always highly significant and positive in both models (Lag model and Durbin Model). This parameter indicates the influence of neighbouring municipalities' cooperation decisions on the cooperation choice of municipalities. We thus provide evidence that cooperation decision-making is largely dependent on the decision of neighbours, independently of the definition of spatial matrix.

Note that this spatial coefficient is decreasing along the different specifications in Table 1 (Spatial Lag Model). LPM results are around 0.5 to 0.7 while CF estimates are always lower, i.e. around 0.1. L-GMM results are less informative, ranging from 0.5 to 0.2 but showing also a 0.7 non significant result when policy change instrument is included together with geo-contiguity matrix. Tables A2 and A3 shows the F-test for the first stage of the CF estimations for each weight matrix. These results were expected since the instrument designed from the policy change is stronger. [16]

In Table 2 (Durbin model), all the spatial coefficients are positive and significant (except those in L-

16. The F-tests reported in Tables A2 and A3 are computed including the instrument and its specific Mundlak-Chamberlain device. Both Tables also show an F-test that is computed excluding the Mundlak-Chamberlain devices, thus showing the power of the instrument alone. We observe that the relevancy of the instrument increases when specified via semi-parametric regression.

GMM when the instrument is specified through policy change). We confirm that the coefficients estimated via CF are lower than LPM in most specifications. Tables A2 and A3 show the F-tests, which are in line with the previous results: the instrument is weaker if specified in the traditional way rather than through a policy change. Note that the only case in which LPM shows a lower result respect to CF is associated with a first stage F-test that is lower than 10 (see Table A6, column 5).

The statistical interpretation of our parameters suggests that both the specification chosen (i.e. LPM, L-GMM or CF) and the validity/relevancy of identification strategy matters in shaping the results.

Within each specification shown in Tables 1 and 2, the spatial lag coefficient increases together with the relevancy of the instrument (from the traditional instrument W^2X to the external instrument WZ used in a semi-parametric way). We believe that this is a consequence of the fact that different instruments identify different local average treatment effect (LATE). [17] Using the CF estimations in which the first stage linear regression is specified against those with a semi-parametric first stage, we find a similar effect: the coefficient increases with the relevancy of the instrument. In this case, even if the instrumental variable is the same and thus compliers are similar, the first stage specification changes the extent of treatment and leads to a different LATE.

We find a negative coefficient for $\hat{\nu}$, the predicted residual computed after the first stage regression. As explained in Baum-Snow and Ferreira (2015), this coefficient has a useful economic interpretation. It is negative for those observations, which were treated more than expected. In our case, localities that receive more state grants than expected are more likely to remain isolated. More specifically, a weighted average increase in unexpected block grant predicts a weighted average decrease in cooperation. Although the direction of the bias correction is different from the one observed in Lyytikainen (2012) and Baskaran (2014), we similarly find that the coefficient is moved in a unique direction (i.e. upward) when the instrument becomes stronger.

Our results are also in line with those of Lyytikainen (2012) and Baskaran (2014) since they show, in the case of a linear spatial model, that employing IV methods associated with a careful choice of instruments can result in clearly different and more reliable estimates of the spatial interaction term. We believe that our results support the previous literature leading to the conclusion that a better specification of the model and an adequate identification strategy ameliorates the assessment of the spatial interaction effects.

From an economic point of view, we can interpret this positive spatial lag coefficient in different ways. First, the territorial contiguity constraint, imposes that two municipalities can cooperate provided that they share a common border. Obviously, a municipality is more likely to cooperate if it is surrounded by cooperating municipalities. We can expect also that information related to cooperation (expected re-

17. In other words the compliers identified by each instrument (W^2X or WZ) are different within each specification.

venues, expected state grant, etc.) will be more easily obtained by an independent municipality that is located next to a group of cooperating municipalities. This is also in line with the literature on yardstick competition (Besley and Case, 1995) and tax competition (Brueckner, 2003). An incumbent may try to avoid stigmatisation and non-reelection by mimicking the cooperation (or isolation) decision of neighbours. We cannot exclude a "tax competition" explanation for this result. Local governments are likely to avoid capital flight from their territories by cooperating in a close community, which provides a new range of public services to firms and households.

To disentangle between these three main sources of spatial interaction (institutional constraint/yardstick competition/tax competition), the existing empirical literature recommends the estimation of auxiliary equations (see e.g. Besley and Case, 1995; Brueckner, 2003). More precisely, identifying specific yardstick competition behaviour would require the estimation of an equation that relates voter approval of an incumbent to taxes in neighbouring jurisdictions. To do so, a more complete dataset on election results would be necessary. Unfortunately, such a dataset exists in France for municipalities with more than 10,000 inhabitants only.

Since the geo-political matrix combines both criteria - contiguity and political proximity - at the same time, we need to disentangle the effect of political proximity weighted by distance from "pure" spatial contiguity on the probability to cooperate. The naive CF approach allows us to address this question by including both matrices in the same specification and considering the two spatially lagged values of the dependent variable (one derived from the contiguity matrix and the other from the geo-political matrix) as two different endogenous variables. The model to be estimated therefore is given by:

$$C_{it} = \phi(\rho_1 \sum_{j=1}^{N} w_{ij} C_{jt} + \rho_2 \sum_{j=1}^{N} w_{ij}^{geoPol} C_{jt} + \beta X_{it}^{'} + \eta_i + e_{it} + \hat{\nu}_{1i,t} + \hat{\nu}_{2i,t}); \qquad (8)$$

where w_{ij} and w_{ij}^{geoPol} are the spatial weights of the contiguity matrix and geo-political matrix respectively, ρ_1 and ρ_2 are the associated spatial parameters and $\hat{\nu}_1$ and $\hat{\nu}_2$ are the residuals of the two first-step regressions of $\sum_{j=1}^{N} w_{ij} C_{jt}$ and $\sum_{j=1}^{N} w_{ij}^{geoPol} C_{jt}$ respectively.

Table 3

The estimation results of Eq. (6) are reported in Table 3. Interestingly, we find that the spatial parameter associated with the contiguity matrix remains significant and positive in all specifications while the spatial coefficient associated with the geo-political matrix takes a very low value (positive or negative), which is sometimes not significant. This suggests that our data weakly support the hypothesis that political proximity increases the probability of cooperation. In other words, the estimates seem to suggest that the main effect at work is an imitation effect of geographic neighbours decisions but that imitation effect much less applies to political belief.

However, the existing literature has often provided some evidence on the role of politics on municipal mergers' decisions (Gordon and Knight, 2009; Weese, 2015; Saarimaa and Tukiainen, 2014; Hyytinen et al., 2014). In the French case, which deals with functional cooperation, municipal officials remain in post after joining an inter-municipal jurisdiction and keep some strategic competences within their municipal territories (i.e. responsibility for local urban services, buildings, nurseries, primary schools, sports facilities). Municipal councils also decide whether the municipality will join the inter-municipal jurisdiction (EPCI), which is governed by a board of delegates elected by municipal councils among their members. Therefore, as opposed to the Finnish case addressed by Hyytinen et al. (2014), functional cooperation leaves some power to the municipal officials both in their locality and within the community. Although we could expect some political alignment within the board of inter-municipal officials, the electoral vulnerability due to the composition of the community board is not as strategic in our case as in the unification case addressed by most articles.

If we now turn to the socio-economic covariates, we observe that the coefficient of fiscal revenues per capita (at time $t-1$) is negative and significant in almost all the specifications. As expected, a high level of fiscal revenues at time $t-1$ reduces the probability of joining a community whereas a financial poor locality may have stronger incentives to set an agreement with neighbouring localities. The obvious financial gains from cooperation (through the extra state transfer to the community) and the possible provision of extra public goods are counterbalanced by each cooperating municipality losing some fiscal autonomy through devolution of competences and tax revenues to the community. The lagged covariate that gives the impact of the neighbours' fiscal revenues on the decision to cooperate is also in line with this explanation and shows, that the municipal fiscal revenues (at time $t-1$) of the neighbouring municipalities may be relevant to the cooperation decision. When this lagged covariate is significant, its coefficient is negative suggesting that, being surrounded by rich municipalities does not encourage cooperation among local governments. Both results are in line with Sorensen (2006), which shows that for Norway high-revenue municipalities prefer not to merge with poorer neighbours.

In relation to the remaining economic and socio-demographic determinants of cooperation, once we control for the cooperation decisions of neighbourhood using the control function approach, we find that most economic and socio-demographic characteristics of the municipalities exhibit the expected sign. We find a positive sign of density in line with the need to reduce free-riding behaviour from citizens who do not live in a city but benefit from her public good provision. In others words, dense municipalities tend to share or to transfer some competences and the associated supply of services, to the inter-municipal group. We often get negative and significant sign for the share of elderly people since this category of the population has comparatively lower needs related to investment spending, which are financed by the inter-municipal community. Also, social services are not legally compulsory for inter-municipal communities. Municipalities with a high proportion of elderly people do not have necessarily the incentive to transfer those competences. For those covariates, we find similar results in terms of neighbourhood characteristics. Municipalities surrounded by dense localities has a higher probability to join an inter-municipal agreement. Those who are located at the proximity of local governments with a high proportion of elderly people cooperate less.

Achieving scale economies and internalizing spillovers effects clearly appears as a motivation for most municipalities to cooperate. However, own economic and demographic characteristics such as a high level of fiscal revenues or a high proportion of elderly people may lead the municipalities to reject cooperation. Our results show that there is a trade-off between remaining isolated - which implies that public services are more tailored to citizens/voters - and the achievement of scale economies or the internalisation of spillover effects within a large community.

6 Concluding remarks

This paper explored empirically the determinants of cooperation decision-making by French municipalities, an interesting setting that involves municipalities decide to join an upper tier of government for the joint provision of some public goods. Unlike merger, this decision does not remove full power from the local governments and local mayors who retain responsibility for certain services.

To disentangle the political, economic and socio-demographic determinants, we use a unique panel data of 1,056 municipalities within a French region (Brittany). We employed a control function approach to develop a binary discrete choice model with spatial interaction. We exploited the exogenous variation in the block grant computation as an exclusion restriction to estimate the decision to cooperate in response to neighbours' decisions.

The estimation results provide evidence that the choice to join an inter-municipal community is largely dependent on the decision of geographic neighbours. Unlike mergers case, we find that French data less support the influence of politics in the cooperation decisions.

This paper is likely to contribute to ongoing debate on the re-organization of sub-national jurisdictions since it promotes the idea that voluntary cooperation among local governments is more difficult to emerge when municipalities have different socio-economic characteristics. In terms of public policies, state can not implement such agreement without allocating extra state grants to improve the incentives to cooperate for the richest local governments.

We contribute also by developing a novel specification of a non-linear spatial regression. We run a complete set of tests for inference and identification, which resolves many computational issues. To our knowledge, this study is the first to use a control function in this context.

Further research should be done to enhance our understanding of the determinants of fiscal cooperation, such as the bargaining power that a municipality may have within the inter-municipal community.

7 References

[1] Alesina, A., Spolaore, E. (1997), On the Number and Size of Nations, *Quarterly Journal of Economics*, 112(4): 1027-1056.

[2] Alesina, A., Spolaore, E. (2003), *The Size of Nations*, Cambridge, MA: MIT Press.

[3] Alesina, A., Spolaore, E. (2005), War, Peace and the Size of Countries, *Journal of Public Economics*, 89(7): 1333-1354.

[4] Alesina, A., Spolaore, E. (2006), Conflict, Defense Spending, and the Number of Nations, *European Economic Review*, 50(1): 91-120.

[5] Aumann, R.J, Myerson, R.B. (1988), Endogenous Formation of Links between Players and of Coalitions: An application of the Shapley Value, *Cambridge University Press*, 175-192.

[6] Baskaran, T. (2014), Indentifying local tax mimicking with administrative borders and a policy reform, *Journal of Public Economics*, 118: 41-51.

[7] Besley T., Case, A. (1995), Incumbent behavior: Vote seeking, tax setting and yardstick competition,*American Economic Review*, 85(1): 25-45.

[8] Bloch, F., Sanchez-Pages, S., Soubeyran, R. (2006), When does universal peace prevail? Secession and group formation in conflict, *Economics of Governance*, 7(1): 3-29.

[9] Blume, L., Blume, T. (2007), The economic effects of local authority mergers: Empirical evidence for German city regions, *Annals of Regional Science*, 41: 689-713.

[10] Bolton, P., Roland, R. (1997), The Breakup of Nations: A Political Economy Analysis, *Quarterly Journal of Economics*, 112(4): 1057-1090.

[11] Borge, L.E, Rattso, J. (2002), Spending Growth With Vertical Fiscal Imbalance: Decentralized Government Spending in Norway, 1880-1990, *Economics and Politics, Wiley Blackwell*,14(3): 208-373.

[12] Brasington, D. (1999), Joint provision of public goods: The consolidation of school districts, *Journal of Public Economics*, 73(3): 373-393.

[13] Brasington, D. (2003a), Size and School District Consolidation: Do Opposites Attract?, *Economica*, 70(280): 673-690.

[14] Brasington, D. (2003b), Snobbery, Racism, or Mutual Distaste: What Promotes and Hinders Cooperation in Local Public-Good Provision?, *Review of Economics and Statistics*, 85: 874-883.

[15] Brueckner, J.K. (2003), Strategic interaction among governments: an overview of empirical studies, *International regional science review*, 26(2): 175-188.

[16] Di Porto, E., Merlin, V., Paty, S. (2013), Cooperation among local governments to deliver public services: A structural bivariate response model with fixed effects and endogenous covariate. Working paper, GATE 2013-04.

[17] de Mello, L., Lago-Penas, S., (2013) Local government cooperation for joint provision: the experiences of Brazil and Spain with inter-municipal consortia, Chapters, in: The Challenge of Local Government Size, chapter 9: 221-241, Edward Elgar.

[18] Dexia (2012), Subnational public finance in the European Union, Dexia.

[19] Elhorst, J.P. (2003), Unconditional maximum likelihood estimation of dynamic models for spatial panels. Research Report 03C27, University of Groningen, Research Institute SOM (Systems, Organisations and Management).

[20] Gibbons, S., Overman, H. (2012), Mostly Pointless Spatial Econometrics?, *Journal of Regional Science*, 52(2): 172-191.

[21] Gibbons, S., Overman H.G., Patacchini, E. (2015), Spatial methods, Handbook of Urban and Regional Economics, Volume 5.

[22] Gordon, N. (2004), Do federal grants boost school spending?, *Journal of Public Economics*, 88(9-10): 1771-92.

[23] Gordon, N., Knight, B. (2009), A spatial merger estimator with an application to school district consolidation, *Journal of Public Economics*, 93(5-6): 752-765.

[24] Hinnerich, T. (2009), Do merging local governments free ride on their counterparts when facing boundary reform?, *Journal of Public Economics*, 93(5-6): 721-728.

[25] Hulst, R., Van Montfort, A., Haveri, A., Airaksinen, J., Kelly, J. (2009), Institutional shifts in intermunicipal service delivery, *Public Organization Review*, 9: 263-285.

[26] Hyytinen, A., Saarimaa, T., Tukiainen, J. (2014), Electoral vulnerability and size of local governments: Evidence from voting on municipal mergers, *Journal of Public Economics*, 120: 193-204.

[27] Klier, T., McMillen, D. (2008), Evolving Agglomeration In The U.S. Auto Supplier Industry, *Journal of Regional Science*, 48(1): 245-267.

[28] Lago-Penas, S., Martinez-Vazquez, J. (2013), The Challenge of Local Government Size: Theoretical Perspectives, International Experience and Policy Reform, Edward Elgar Pub.

[29] LeSage, J. (2000), Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models, *Geographical Analysis*, 32(1): 19-35.

[30] Lyytikainen, T. (2012), Tax competition among local governments: Evidence from a property tax reform in Finland, *Journal of Public Economics*, 96 (7-8): 584-595.

[31] Manski, C.F. (1993), Identification of Treatment Response with Social Interactions, *Econometrics Journal*, 16(1): 1?23.

[32] Myerson, R. (1977), Graphs and Cooperation in Games. *Math. Oper. Res.*, 2: 225-229.

[33] Oates, W.E. (1972), Fiscal Federalism, *Journal of Economic Issues*, 6(4): 225-227.

[34] Papke, E., Wooldridge, J. (2008), Panel data methods for fractional response variables with an application to test pass rates, *Journal of Econometrics*, 145(1-2): 121-133.

[35] Papke, E. (2005), The effects of spending on test pass rates: evidence from Michigan, *Journal of Public Economics*, 89(5-6): 821-839.

[36] Pinkse, J., Slade, M.E. (1998), Contracting in space: An application of spatial statistics to discrete-choice models, *Journal of Econometrics*, 85(1): 125-154.

[37] Poel, D.H. (2000), Amalgamation Perspectives: Citizen Responses to municipal Consolidation, *Canadian Journal of Regional Science*, 23(1): 31-48.

[38] Reingewertz, Y. (2012), Do municipal amalgamations work? Evidence from municipalities in Israel, *Journal of Urban Economics*, 72(2): 240-251.

[39] Saarimaa, T., Tukiainen, J. (2014), I Don't Care to Belong to Any Club That Will Have Me as a Member: Empirical Analysis of Municipal Mergers. *Political Science Research and Methods*, 2: 97-117.

[40] Salmon, P. (1987), Decentralization as an incentive scheme,*Oxford review of economic policy* 3(2): 24-43.

[41] Shenoy, L. (1976), On coalition formation: A game theoretical approach, *International Journal of Game Theory*, 8: 133-164.

[42] Sorensen, R.J. (2006), Local government consolidation: The impact of political transaction costs, *Public Choice*, 127: 75-96.

[43] Spolaore, E. (2004), Is small really so ugly?, *World trade Review*, 30(3): 447-452.

[44] Tiebout, C.M. (1956), A Pure Theory of Local Expenditures, *Journal of Political Economy*, 64(5): 416-424.

[45] Weese, E. (2015), Political Mergers as Coalition Formation: An Analysis of the Heisei Municipal Amalgamations. *Quantitative Economics*, 6(2): 257-307.

[46] Won, J.H., Lim, J., Kim, S.J., Rajaratnam, B. (2013), Condition-number-regularized covariance estimation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(3): 427-450.

8 Appendix

A - Bootstrap procedure

Due to the two-step regression in the CF approach, the standard errors in the second stage must be adjusted for the first-stage estimation. As Papke and Wooldridge (2008) suggest, we therefore use a bootstrap procedure.

Given the spatial structure of the data, ignoring the spatial correlation in the bootstrap procedure and directly resampling the observations will destroy the correlation structure among the original observations. To improve the bootstrap method we follow Solow (1985) who introduces a particular bootstrap procedure for spatially correlated data. The basic idea involves transforming correlated observations to uncorrelated ones, creating a bootstrapping sample from the latter, and transforming back to a quasibootstrap sample from the original data.

We extend the Solow's procedure to panel data, by allowing block-bootstrap for the entire time series and using a wild bootstrap to take into account the presence of heteroskedasticity.

Given the model $C_{it} = \hat{\phi}(\cdot) + \hat{e}_{it}$ for $i = 1, ..., N$ and $t = 1, ..., T$, where C_{it} is the dependent variable, $\hat{\phi}(\cdot)$ is the estimate for the deterministic component and $\hat{e}_{it}$ are the residuals, the bootstrap procedure consists of the following steps.

1. Estimate the covariance matrix Σ. [18]
2. Decompose the resulting positive definite covariance matrix $\hat{\Sigma}$ using the Cholesky decomposition $\hat{\Sigma} = \hat{L}\hat{L}^T$, where $\hat{L}$ is a lower triangular $NT \times NT$ matrix.
3. Use the Cholesky decomposition matrix inverse, $\hat{L}^{-1}$, to decorrelate the spatial errors, i.e., $\hat{\epsilon} \equiv \hat{L}^{-1}\hat{e}$.
4. Centers the $\hat{\epsilon}$ and obtain $\tilde{\epsilon} = \hat{\epsilon} - \bar{\hat{\epsilon}}$.
5. Generate B independent block wild-bootstrap samples in two-steps:
   (a) draw with replacement NT residuals from the following two-point distribution:

$$\tilde{\epsilon}^{\star} = \begin{cases} (1-\sqrt{5})/2 & \text{with } p = (5+\sqrt{5})/10 \\ (1+\sqrt{5})/2 & \text{with } 1-p \end{cases}$$

   keeping the temporal structure intact;

   (b) transforming to recorrelated the bootstrap residuals to form the quasibootstrap sample $C^{\star}_{it} = \hat{\phi}(\cdot) + \hat{L}\tilde{\epsilon}^{\star}$ and calculate $\sum_{j=1}^{N} w_{ij} C^{\star}_{it}$.

18. Given the high dimensionality of our covariance matrix, we use a maximum likelihood covariance estimator with condition number constraint to get an invertible and well-conditioned matrix (see Won *et al.*, 2013). This estimate is carried out using the "CondReg" R package.

6. Estimate the model for each $C_1^\star, ..., C_B^\star$ and take the estimated parameters of the second stage regression.
7. Compute for each bootstrap sample, $b = 1, ..., B$, and for each explanatory variable, $k = 1, ..., K$, the average partial effect APE_b^k.
8. Calculate the bootstrap standard errors and the p-values.

In our estimates we set $B = 1000$.

B - Descriptive Statistics

Table A1

C - Estimation results

Table A2

Table A3

Table A4

Table A5

Table A6

Table A7

9 Estimates

TABLE 1 – LPM, L-GMM and CF Estimations of ρ in the Spatial Lag Model

Lag Model						
	Contiguity Matrix			**Geo(Contiguity)-Political Matrix**		
	Instruments: W^2X			*Instruments:* W^2X		
	LPM	L-GMM	CF	LPM	L-GMM	CF
ρ	0.707***	0.495***	0.170***	0.534***	0.227***	0.073***
$\hat{\nu}$			0.305***			0.365***
	Instruments: WZ			*Instruments:* WZ		
	LPM	L-GMM	CF	LPM	L-GMM	CF
ρ	0.707***	0.735	0.202***	0.534***	0.451*	0.071***
$\hat{\nu}$			0.271***			0.264***
	Instruments: $s(WZ)$			*Instruments:* $s(WZ)$		
	LPM	L-GMM	CF	LPM	L-GMM	CF
ρ	0.707***	-	0.559***	0.534***	-	0.133***
$\hat{\nu}$			-0.094***			0.247***

Note. For CF Average Partial Effects (APEs) are reported.
Significance levels: *p<0.1; **p<0.05; ***p<0.01

TABLE 2 – LPM, L-GMM and CF Estimations of ρ in the Spatial Durbin Model

Durbin Model						
	Contiguity matrix			**Geo(Contiguity)-Political Matrix**		
	Instruments: W^2X			*Instruments:* W^2X		
	LPM	L-GMM	CF	LPM	L-GMM	CF
ρ	0.706***	0.437***	0.296***	0.532***	0.248***	0.302***
$\hat{\nu}$			0.182***			0.131***
	Instruments: WZ			*Instruments:* WZ		
	LPM	L-GMM	CF	LPM	L-GMM	CF
ρ	0.706***	0.428	0.299***	0.532***	0.388	0.650***
$\hat{\nu}$			0.179***			-0.221***
	Instruments: $s(WZ)$			*Instruments:* $s(WZ)$		
	LPM	L-GMM	CF	LPM	L-GMM	CF
ρ	0.706***	-	0.647***	0.532***	-	0.507***
$\hat{\nu}$			-0.177***			-0.083***

Note. For CF Average Partial Effects (APEs) are reported.
Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table 3 – CF Estimations of ρ_1 and ρ_2 in the Lag and Spatial Durbin Models

	Lag Model	***Durbin Model***
	Instruments: W^2X	
	CF	CF
ρ_1	0.207***	0.302***
ρ_2	-0.015***	0.005*
$\hat{\nu}_1$	0.239***	0.160***
$\hat{\nu}_2$	0.050***	0.015**
	Instruments: WZ	
	CF	CF
ρ_1	0.122***	0.281***
ρ_2	0.087***	-0.003
$\hat{\nu}_1$	0.344***	0.176***
$\hat{\nu}_2$	-0.077***	0.027***
	Instruments: $s(WZ)$	
	CF	CF
ρ_1	0.571***	0.640***
ρ_2	-0.012***	0.009***
$\hat{\nu}_1$	-0.122***	-0.186***
$\hat{\nu}_2$	0.030***	0.009*

Note. Average Partial Effects (APEs) are reported.
Significance levels: *p<0.1; **p<0.05; ***p<0.01

A Tables

	Density_{t-1}	Elderly_{t-1}	Revenues_pc_{t-1}	WDensity_{t-1}	WElderly_{t-1}	$\text{WRevenues\_pc}_{t-1}$
min	8.77	0.06	0.06	87.19	0.18	0.18
max	35281.29	0.53	7.01	4784.50	0.30	0.40
median	306.56	0.25	0.21	762.36	0.25	0.25
mean	753.59	0.25	0.25	752.25	0.25	0.25
std.dev	1928.14	0.08	0.21	491.14	0.03	0.03

Table A.1 – Descriptive statistics

Table A.2 – Results of the First-stage Regressions for CF with Contiguity Spatial Matrix.

	Model Lag Contiguity			*Model Durbin Contiguity*		
Instruments:	W^2X	WZ	$s(WZ)$	W^2X	WZ	$s(WZ)$
Time FE:	YES	YES	YES	YES	YES	YES
$\text{W}^2\text{Density}_{t-1}$	0.000***	–	–	0.000***		
	(0.000)	–	–	(0.000)		
$\text{W}^2\text{Elderly}_{t-1}$	12.270***	–	–	12.350***		
	(1.301)	–	–	(2.020)		
$\text{W}^2\text{Revenues\_pc}_{t-1}$	−0.688***	–	–	−0.467		
	(0.235)	–	–	(0.371)		
$\overline{\text{W}^2\text{Density}}_{t-1}$	0.000	–	–	0.000**		
	(0.000)	–	–	(0.000)		
$\overline{\text{W}^2\text{Elderly}}_{t-1}$	−11.670***	–	–	−11.450***		
	(1.300)	–	–	(2.020)		
$\overline{\text{W}^2\text{Revenues\_pc}}_{t-1}$	0.610**	–	–	0.671*		
	(0.240)	–	–	(0.372)		
WZ_{it}	–	0.000	–	–	0.000	–
	–	(0.000)	–	–	(0.000)	–
$\overline{WZ_{it}}$	–	0.001***	–	–	0.001***	–
	–	(0.000)	–	–	(0.000)	–
$s(WZ_{it}, \overline{WZ_{it}})$	–	–	24.69***(a)	–	–	24.84***(a)
F-test on instruments(b)	24.256***	47.657***	24.294***	12.045***	35.260***	23.644***
F-test on instruments(c)	32.933***	0.0385	35.817***	14.382***	0.2931	27.050***
$\bar{R}^2$	0.12	0.11	0.17	0.12	0.12	0.18
N.Obs	7392	7392	7392	7392	7392	7392

Note. Standard errors in parenthesis. Significance levels: *p<0.1; **p<0.05; ***p<0.01.
(a) Estimated Degrees of Freedom (edf) for semiparametric model.
(b) F-test on instruments *including* Mundlak.
(c) F-test on instruments *excluding* Mundlak.
Coefficients of exogenous regressors are not reported for sake of shortness.

Table A.3 – Results of the First-stage Regressions for CF with Geo(Contiguity)-Political Spatial Matrix.

	Model Lag Contiguity			*Model Durbin Contiguity*		
Instruments:	W^2X	WZ	$s(WZ)$	W^2X	WZ	$s(WZ)$
Time FE:	YES	YES	YES	YES	YES	YES
$\text{W}^2\text{Density}_{t-1}$	0.000**	–	–	0.000		
	(0.000)	–	–	(0.000)		
$\text{W}^2\text{Elderly}_{t-1}$	9.390***	–	–	8.249***		
	(1.139)	–	–	(1.387)		
$\text{W}^2\text{Revenues\_pc}_{t-1}$	0.081	–	–	0.288		
	(0.227)	–	–	(0.273)		
$\overline{\text{W}^2\text{Density}}_{t-1}$	0.000*	–	–	0.000		
	(0.000)	–	–	(0.000)		
$\overline{\text{W}^2\text{Elderly}}_{t-1}$	−7.300***	–	–	−5.797***		
	(1.138)	–	–	(1.389)		
$\overline{\text{W}^2\text{Density\_revenues\_pc}}_{t-1}$	0.321	–	–	0.196		
	(0.230)	–	–	(0.274)		
WZ_{it}	–	0.000	–	–	0.000	–
	–	(0.000)	–	–	(0.000)	–
$\overline{WZ_{it}}$	–	0.001***	–	–	0.00*	–
	–	(0.000)	–	–	(0.000)	–
$s(WZ_{it}, \overline{WZ_{it}})$	–	–	28.47***(a)	–	–	27.21***(a)
F-test on instruments(b)	634.07***	61.447***	61.451***	103.14***	1.5915	23.579***
F-test on instruments(c)	24.185***	0.7019	65.205***	13.052***	1.0874	21.444***
$\bar{R}^2$	0.39	0.10	0.26	0.40	0.12	0.40
N.Obs	7392	7392	7392	7392	7392	7392

Note. Standard errors in parenthesis. Significance levels: *p<0.1; **p<0.05; ***p<0.01.
(a) Estimated Degrees of Freedom (edf) for semiparametric model.
(b) F-test on instruments *including* Mundlak.
(c) F-test on instruments *excluding* Mundlak.
Coefficients of exogenous regressors are not reported for sake of shortness.

Table A.4 – Results of the Second-stage Regression for CF Estimations with Contiguity Spatial Matrix.

	Model Lag Contiguity			*Model Durbin Contiguity*		
Instruments:	W^2X	WZ	$s(WZ)$	W^2X	WZ	$s(WZ)$
Mundlak:	YES	YES	YES	YES	YES	YES
Time FE:	YES	YES	YES	YES	YES	YES
Constant	0.020*	−0.011	−0.331***	−0.066***	−0.065***	−0.369***
	(0.012)	(0.024)	(0.009)	(0.022)	(0.027)	(0.010)
ρ	0.170***	0.202***	0.559***	0.296***	0.299***	0.647***
	(0.015)	(0.028)	(0.011)	(0.026)	(0.031)	(0.011)
$\hat{\nu}$	0.305***	0.271***	−0.094***	0.182***	0.179***	−0.177***
	(0.015)	(0.029)	(0.011)	(0.025)	(0.031)	(0.011)
Density_{t-1}	0.000***	0.000***	0.000**	0.000***	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Elderly_{t-1}	0.159**	0.079	−0.501***	0.112*	0.125*	−0.196*
	(0.076)	(0.087)	(0.076)	(0.080)	(0.086)	(0.092)
Revenues_pc_{t-1}	−0.070***	−0.073***	−0.043***	−0.041***	−0.043***	−0.032***
	(0.008)	(0.009)	(0.010)	(0.008)	(0.008)	(0.008)
WDensity_{t-1}	–	–	–	0.000***	0.000***	0.000**
	–	–	–	(0.000)	(0.000)	(0.000)
WElderly_{t-1}	–	–	–	0.166	0.078	−1.857***
	–	–	–	(0.202)	(0.206)	(0.151)
$\text{WRevenues\_pc}_{t-1}$	–	–	–	−0.183***	−0.181***	−0.070***
	–	–	–	(0.018)	(0.020)	(0.019)
$\bar{R}^2$	0.53	0.53	0.53	0.53	0.53	0.53
AIC_c	3421.92	3430.29	3433.80	3375.92	3376.41	3370.90
N.Obs	7392	7392	7392	7392	7392	7392

Note. Average Partial Effects (APEs) and wild bootstrap standard errors (in parenthesis). Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table A.5 – Results of LPM-IV Estimations with Contiguity Spatial Matrix.

	Model Lag Contiguity			*Model Durbin Contiguity*		
Instruments:	W^2X	WZ	$s(WZ)$	W^2X	WZ	$s(WZ)$
Mundlak:	YES	YES	YES	YES	YES	YES
Time FE:	YES	YES	YES	YES	YES	YES
$\hat{\rho}$	0.894***	1.917	1.298***	0.829***	1.414*	1.406***
	(0.123)	(1.976)	(0.278)	(0.183)	(0.798)	(0.355)
Density_{t-1}	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Elderly_{t-1}	–0.090	–1.356	–0.508	–0.003	–0.537	–0.464
	(0.362)	(2.468)	(0.464)	(0.390)	(0.802)	(0.467)
Revenues_pc_{t-1}	0.027	0.008	0.022	0.032	0.025	0.025
	(0.054)	(0.066)	(0.054)	(0.057)	(0.058)	(0.057)
WDensity_{t-1}	–	–	–	0.000	0.000	0.000
	–	–	–	(0.000)	(0.000)	(0.000)
WElderly_{t-1}	–	–	–	0.218	–1.605	–1.362
	–	–	–	(0.840)	(2.537)	(1.219)
$\text{WRevenues\_pc}_{t-1}$	–	–	–	–0.023	–0.030	–0.031
	–	–	–	(0.109)	(0.109)	(0.109)
N.Obs	7392	7392	7392	7392	7392	7392

Note. Standard errors in parenthesis. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table A.6 – Results Second-stage regression for CF Estimations with Geo(Contiguity)-Political Matrix.

	Model Lag Geo(Contiguity)-Political			*Model Durbin Geo(Contiguity)-Political*		
Instruments:	W^2X	WZ	$s(WZ)$	W^2X	WZ	$s(WZ)$
Mundlak:	YES	YES	YES	YES	YES	YES
Time FE:	YES	YES	YES	YES	YES	YES
Constant	0.176***	0.166***	0.109***	0.051***	−0.144***	−0.061***
	(0.004)	(0.023)	(0.007)	(0.009)	(0.042)	(0.008)
ρ	0.073***	0.071***	0.133***	0.302***	0.650***	0.507***
	(0.003)	(0.025)	(0.006)	(0.013)	(0.068)	(0.012)
$\hat{\nu}$	0.365***	0.264***	0.247***	0.131***	−0.221***	−0.083***
	(0.004)	(0.026)	(0.007)	(0.013)	(0.068)	(0.011)
Density_{t-1}	0.000***	0.000***	0.000**	0.000***	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Elderly_{t-1}	0.500**	0.764***	0.6181***	0.397*	0.264***	0.362***
	(0.103)	(0.115)	(0.109)	(0.108)	(0.107)	(0.114)
Revenues_pc_{t-1}	−0.110***	−0.128***	−0.116***	−0.091***	−0.082***	−0.088***
	(0.024)	(0.018)	(0.021)	(0.026)	(0.026)	(0.029)
WDensity_{t-1}	–	–	–	0.000	0.000	0.000
	–	–	–	(0.000)	(0.000)	(0.000)
WElderly_{t-1}	–	–	–	0.083	−1.681***	−0.976***
	–	–	–	(0.165)	(0.319)	(0.163)
$\text{WRevenues\_pc}_{t-1}$	–	–	–	−0.068**	−0.043	−0.056*
	–	–	–	(0.034)	(0.034)	(0.036)
$\bar{R}^2$	0.38	0.28	0.31	0.38	0.38	0.38
AIC_c	4335.64	4820.28	4738.82	4309.29	4318.07	4314.76
N.Obs	7392	7392	7392	7392	7392	7392

Note. Average Partial Effects (APEs) and wild bootstrap standard errors (in parenthesis). Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table A.7 – Results of Second-stage regression for CF Estimations with Two Spatial Matrices.

	Model Lag Two Matrices			*Model Durbin Two Matrices*		
Instruments:	W^2X	WZ	$s(WZ)$	W^2X	WZ	$s(WZ)$
Mundlak:	YES	YES	YES	YES	YES	YES
Time FE:	YES	YES	YES	YES	YES	YES
Constant	0.004	−0.020	−0.331***	−0.077***	−0.052**	−0.372***
	(0.012)	(0.024)	(0.009)	(0.022)	(0.027)	(0.010)
ρ_1	0.207***	0.122***	0.571***	0.302***	0.281***	0.640***
	(0.014)	(0.042)	(0.012)	(0.026)	(0.032)	(0.011)
ρ_2	−0.015***	0.087***	−0.012***	0.005*	−0.003	0.009***
	(0.003)	(0.025)	(0.005)	(0.003)	(0.003)	(0.003)
$\hat{\nu}_1$	0.239***	0.344***	−0.122***	0.160***	0.176***	−0.186***
	(0.015)	(0.043)	(0.012)	(0.026)	(0.031)	(0.012)
$\hat{\nu}_2$	0.050***	−0.077***	0.030***	0.015**	0.027***	0.009*
	(0.006)	(0.026)	(0.005)	(0.007)	(0.008)	(0.006)
Density_{t-1}	0.000***	0.000***	0.000**	0.000***	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Elderly_{t-1}	0.112*	0.102	−0.508***	0.110*	0.153**	−0.190**
	(0.076)	(0.086)	(0.075)	(0.080)	(0.086)	(0.092)
Revenues_pc_{t-1}	−0.070***	−0.079***	−0.043***	−0.042***	−0.045***	−0.033***
	(0.008)	(0.009)	(0.011)	(0.008)	(0.008)	(0.009)
WDensity_{t-1}	–	–	–	0.000***	0.000***	0.000*
	–	–	–	(0.000)	(0.000)	(0.000)
WElderly_{t-1}	–	–	–	0.104	0.171	−1.858***
	–	–	–	(0.203)	(0.207)	(0.151)
$\text{WRevenues\_pc}_{t-1}$	–	–	–	−0.180***	−0.186***	−0.072***
	–	–	–	(0.018)	(0.021)	(0.019)
$\bar{R}^2$	0.53	0.53	0.53	0.53	0.53	0.53
AIC_c	3418.01	3432.70	3435.07	3378.28	3377.86	3373.34
N.Obs	7392	7392	7392	7392	7392	7392

Note. Average Partial Effects (APEs) and wild bootstrap standard errors (in parenthesis). Significance levels: *p<0.1; **p<0.05; ***p<0.01