

CIRJE-F-951

**Information Sharing, Neighborhood Demarcation,
and Yardstick Competition: An Empirical Analysis of
Intergovernmental Expenditure Interaction in Japan**

Masayoshi Hayashi
The University of Tokyo

Wataru Yamamoto
Graduate School of Economics, The University of Tokyo

December 2014; Revised in June 2016

CIRJE Discussion Papers can be downloaded without charge from:

<http://www.cirje.e.u-tokyo.ac.jp/research/03research02dp.html>

Discussion Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Discussion Papers may not be reproduced or distributed without the written consent of the author.

Information Sharing, Neighborhood Demarcation, and Yardstick Competition: An Empirical Analysis of Intergovernmental Expenditure Interaction in Japan

Masayoshi Hayashi^a
Wataru Yamamoto^b

Abstract

The Japanese government provides information on local fiscal performance through the *Fiscal Index Tables for Similar Municipalities* (FITS-M). The FITS-M categorize municipalities into groups of “similar localities” and provide them with the fiscal indices of their group members, enabling municipalities to use the tables to identify their “neighbors” (i.e., those in the same FITS-M group) and refer to their fiscal information as a “yardstick” for fiscal planning. We take advantage of this system to estimate municipal spending function. In particular, we examine if the FITS-M help identify a defensible spatial weights matrix that properly describes municipal spending interactions. Our analysis shows that they do. In particular, geographical proximity is significant only between a pair of municipalities within a given FITS-M group, and it does not affect competition between pairs belonging to different groups even if they are located close to each other. This would suggest that the FITS-M work as intended, indicating that spending interaction among Japanese municipalities originates from yardstick competition and not from other types of fiscal competition.

Keywords: *fiscal interaction, local government expenditure, yardstick competition, fiscal federalism, Japan*

JEL Classification: H72, H73, H77

^a Mailing Address: Faculty & Graduate School of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; E-mail: hayashim@e.u-tokyo.ac.jp; Tel: +81-3-5841-5513 (DI); Fax: +81-3-5841-5521.

^b Doctoral Program, Graduate School of Economics, University of Tokyo, Japan.

1. Introduction

Public finance literature has a long tradition of exploring determinants of government, mainly subnational (i.e., local or state), spending (Facchini 2014). In the last decade, an increasing number of studies focus on the roles of *other* governments in explaining government expenditures. To estimate such impacts of other governments, the literature utilizes a spatial weights matrix \mathbf{W} that specifies the group of other governments that affect spending by a given government and the degree of their impacts. In other words, \mathbf{W} defines “neighbors” and reflects spending interdependence among local governments. The literature on spatial econometrics has long recognized that different choices of \mathbf{W} may change the estimation results, and its misspecification would cause flaws in the estimator (Stetzer 1982, Griffith and Lagona 1998, Stakhovych and Bijmolt 2008, Smith 2009). However, it is not always straightforward to find a form of \mathbf{W} that is defensible (Harris et al. 2011, Lundberg 2014). Some, therefore, propose a method that estimates spatial models *without* \mathbf{W} (Folmer and Oud 2008), while others suggest a method that estimates \mathbf{W} as *parameters* by imposing restrictions on it (Beenstock and Felsenstein 2012, Bhattacharjee and Jensen-Butler 2013). While these lines of inquiry constitute viable avenues for studying \mathbf{W} , we may also defend a specific structure of \mathbf{W} by closely looking at institutional mechanisms of intergovernmental relations.

In this paper, we aim to contribute to the literature by constructing a defensible specification of \mathbf{W} that takes advantage of an institutional system of information sharing among local governments, and we use it to distinguish a type of fiscal competition from other types. We base our analysis on the models of “yardstick competition.” Originally coined for competitions in regulated markets (Shleifer 1985), the fiscal federalism literature uses the term to characterize competition based on performance comparison among local governments (Salmon 1987). Drawing from the theory of the principal–agent, the theoretical analyses show that yardstick competition makes local governments respond to other governments’ policies over expenditures (Bivand and Szymanski 1997, Caldeira 2012) as well as taxes (Besley and Case 1995, Wrede 2001, Boarnet et al. 2004). Yardstick competition can be “from the bottom” or “from the top” (Caldeira 2012). If it is from the bottom, which is more typical in the literature, the voters and local policy makers (politicians) constitute the principal and the agent, respectively. The competition originates in the notion that local residents compare the performance of local politicians

in office to those in other jurisdictions, and vote against the incumbent if they perform worse than the latter. On the other hand, if yardstick competition is “from the top,” the central government (principal) evaluates the performance of localities (agents) and rewards those with better performance. While the top-down process was initially intended to characterize non-democracies (Caldeira 2012), an analogous model may be applicable to democracies as well. Examples include cases where benevolent politicians (principal) assess the performance of self-interested bureaucrats (agents) by comparing their performance with that of other bureaucrats in other jurisdictions (Bivand and Szymanski 1997, Revelli and Tovmo 2007).

Whether yardstick competition is from the top or bottom, an important point is that information on other governments matters. Meanwhile, a number of studies on intergovernmental interaction utilize geographical proximity to specify \mathbf{W} , implicitly assuming that more distant entities have less influence. While this assumption may be plausible if the underlying interaction depends on geographical proximity, it may not be so for yardstick competition if geography does not matter in acquiring information on other governments. For example, if the central government freely disseminates fiscal information on local governments, it could reduce the effect of geographical distance on yardstick competition (Revelli 2006).

Another important point is that, in the face of free acquisition of fiscal information, local governments need to identify “neighbors” they compete with. In this regard, the literature argues that local governments compete with those they consider sufficiently “similar” (Case et al. 1993). However, it may not be straightforward for them to pick up relevant “neighbors” if the number of localities is too large, say, more than a thousand, like in some countries. The central government could then help such localities by demarcating a group of localities as their “neighbors,” whose fiscal information serves as a yardstick. Of course, such demarcation specifies the structure of \mathbf{W} in the estimation of government spending interaction.

In an environment where the central government disseminates fiscal information on localities and demarcates the “neighborhood of competition,” local governments may compete with neighbors demarcated by the center. If such a demarcation is independent of geographical proximity, geography should not matter in fiscal interaction based on yardstick competition. Indeed, the Japanese system of the *Fiscal Index Tables for Similar Municipalities* (FITS-M) conveniently approximates such an environment of

information sharing and neighborhood demarcation. The system categorizes localities into groups of *ruiji dantai* or “similar localities” and provides a set of within-group fiscal indices against which localities in a given group can compare their own indices. In fact, Japan’s central government has set up the FITS-M with the clear intention of letting local governments utilize them as *monosashi* or “yardstick” for their fiscal planning (Negishi 2007). Indeed, anecdotal evidence suggests that localities do take note of others in their groups during their annual budgeting process (Nishihama 2007, Matsuki 2010). The FITS-M thus intend to allow local governments to identify “neighbors” within their FITS-M groups and to use their fiscal information as a yardstick.

In this paper, we take advantage of the grouping in the FITS-M to specify \mathbf{W} , estimate a municipal spending function, and compare it with functions estimated using alternative forms of \mathbf{W} based on geographical proximity. If the FITS-M-based \mathbf{W} performs better than the geography-based \mathbf{W} , we may argue that the FITS-M induce spending interaction among Japanese municipalities based on yardstick competition. We develop our ensuing analysis as follows. Section 2 sets the framework for our analysis by discussing the three models of fiscal competition and relating them with popular types of spatial weights. It also introduces the FITS-M-based spatial weights used in this study. Section 3 presents the development of our estimation procedure by discussing the regression model, choice of regressors, data set and its sources, and choice of spatial weights matrices. Section 4 provides the estimation results and discusses their relevance. Finally, Section 5 concludes this study.

2. Spatial Weights and Models of Fiscal Competition

2.1. Reaction function and spatial weights

Assume that there are N localities. Theories of expenditure competition imply that spending by government i , g_i , depends on spending of the other $N - 1$ governments, $\mathbf{g}_{-i} = [g_1, g_2, \dots, g_{i-1}, g_{i+1}, \dots, g_N]'$. We could then express the spending function for i as

$$g_i = f(\mathbf{g}_{-i}, \mathbf{x}_i), \quad (1)$$

where \mathbf{x}_i is a vector of observed factors that affect g_i other than \mathbf{g}_{-i} . When estimating Eq. (1), we may specify it as a linear function

$$g_i = \alpha + \sum_{j \neq i}^N \beta_j \cdot g_j + \sum_{k=1}^K \gamma_k \cdot x_k + u_i, \quad (2)$$

where u_i is the unobserved element that affects g_i , and the Greek letters are the parameters to be estimated. Of course, we cannot easily estimate Eq. (2), since the number of β_j s, the marginal effect of g_j on g_i , is $N - 1$.¹ While recent studies attempt to estimate β_j by imposing restrictions on the model when panel data are available,² the standard trick to cope with this difficulty is to assume that the impact of j 's spending on i 's (β_j) is uniformly proportional to *observable* weights w_{ij} for $\forall j \neq i$ such that

$$\beta_j = w_{ij} \cdot \rho. \quad (3)$$

The literature calls w_{ij} “spatial weight,” although it does not have to be geographically “spatial.” Eq. (3) converts Eq. (2) into a more tractable form, namely

$$g_i = \alpha + \rho \cdot \left(\sum_{j=1}^N w_{ij} \cdot g_j \right) + \sum_{k=1}^K \gamma_k \cdot x_k + u_i, \quad (4)$$

which, with ρ taking a common value, collapse all fiscal interactions across governments into a single weighted variable, $\sum_j w_{ij} g_j$. We can then estimate ρ to assess the individual effects on g_i of each element of \mathbf{g}_{-i} (i.e., β_j s) without directly estimating them.

The spatial weights w_{ij} s comprise the spatial weights matrix \mathbf{W} , which is an $N \times N$ matrix whose $[i, j]$ -th element is w_{ij} and diagonal elements (w_{ii}) are zeros.³ The spatial weights are in turn defined by the “un-normalized” weight W_{ij} , which yields w_{ij} by row normalization such that $w_{ij} \equiv W_{ij}/(\sum_j W_{ij})$. Evidently, \mathbf{W} defines the structure of fiscal interaction among local governments and demarcates “the neighborhood of competition” within which local governments compete against each other exclusively. The structure of \mathbf{W} thus reflects the properties of fiscal competition under investigation, which is likely to differ from one model of competition to another. The choice of \mathbf{W} is therefore important in the estimation of the reaction function.

2.2. Models of fiscal competition and geographical proximity

The interdependence of local government spending can draw on three alternative

¹ This is in contrast to the case where g is spending on a pure public good, where $\beta = \beta_j \forall j \neq i$, which reduces $\sum_{j \neq i} \beta_j g_j$ to $\beta(\sum_{j \neq i} g_j)$ (see Sandler 1992).

² With panel data, we may identify $\beta_j \forall j \neq i$ by imposing restrictions on a spatial weights matrix. Such restrictions include symmetry, where $w_{ij} = w_{ji}$ (Bhattacharjee and Jensen-Butler 2013), and sparsity, where each unit is affected by a limited number of other units only (Ahrens and Bhattacharjee 2015, Bailey et al. 2015). Meanwhile, Bhattacharjee and Holly (2013) consider a specific case that allows us to utilize moment conditions to identify spatial interaction. However, we cannot always justify the use of these restrictions or identifying assumptions (we are plausibly unable to do this for the current case too).

³ The i -th row of \mathbf{W} therefore constitutes a set of spatial weights for the i -th local government.

theoretical arguments. First, it could originate from benefit spillovers of local public goods beyond jurisdictional borders (Williams 1966). If the benefit of spending in one locality spills over to another locality, it affects the welfare of residents in the latter locality. Therefore, when resident welfare matters in local spending decisions, these decisions in one locality affect analogous decisions in others. Second, the fiscal dependence may also emanate from flows of resources (i.e., labor and capital) across jurisdictions (Boadway 1983, Wildasin 1988, Brueckner 2003). If the resources move into localities with more favorable environments, changes in local spending cause them to relocate across jurisdictions. Since individuals (labor) and firms (capital) pay local tax bases, their relocations affect local budgets, allowing spending in a given locality to affect spending in others. Third, the dependence may evolve from yardstick competition based on performance comparison among local governments (Salmon 1987, Case et al. 1993, Besley and Case 1995). Drawing on the principal–agent theory, theoretical analysis shows that yardstick competition makes local governments emulate policies of other governments over expenditures (Bivand and Szymanski 1997, Caldeira 2012) and taxes (Besley and Case 1995).

Geographical proximity would be an important factor in all these three models of fiscal competition. First, geography matters for the spillover model. If the benefits “literally” spill-in from other municipalities, a longer distance may reduce the impacts of the spilt-in benefits. For example, the benefit of rescue service decreases with distance since turnout time is a critical factor (Hanes 2002). If residents move to the municipality of origin to consume benefits like library services (Finney and Yoon 2003), a longer distance increases transportation costs, discouraging their consumption. Second, proximity should also matter for the resource–flow model. When individuals and/or firms move into regions that offer the most favorable fiscal incentives, mobility costs are important and plausibly increasing with distance.⁴ Since a longer distance between a given pair of municipalities discourages flows of resources between them, fiscal interaction originating from resource flows should be more tenuous for a distant pair of localities. Third, geographical proximity may also be important in yardstick competition. Nearby jurisdictions may matter more than distant jurisdictions since the information on the former is likely to spill in more easily than that on the latter (Revelli 2006, Revelli

⁴ For this argument, we discuss movements of existing entities across borders and exclude cases where, say, new firms are born in a given location.

and Tovmo 2007).

Therefore, typical structures of \mathbf{W} in the literature draw on geographical proximity such that the value of w_{ij} decreases as the distance between i and j increases. However, while geography may be the prime factor for fiscal competition originating in benefit spillovers or resource flows, it may not be the only factor for fiscal interaction based on yardstick competition. What really matters in yardstick competition is the information on other governments against which a locality compares its fiscal performance. Geographical closeness is pertinent to the extent that it helps acquire information. In other words, if the acquisition of information is costless and independent of geography, geography should be irrelevant in yardstick competition. Revelli (2006) indeed shows that this is the case in his study on the Social Services Performance Rating (SSPR) in the UK, a nationwide system that evaluates the performance of localities in providing social services and disseminates the evaluation results to all localities. Revelli finds that this system made geographical proximity less relevant in fiscal interaction, since information is now easily obtainable on a nationwide basis.

2.3. Demarcation of neighborhoods and the FITS-M system in Japan

In yardstick competition, local governments are supposed to find a benchmark against which to compare themselves. However, the question of how localities find their benchmark or “neighbors” when geography does not matter remains unanswered. A plausible answer to this question is that localities may take as a benchmark those they consider sufficiently “similar” to themselves (Case et al. 1993). Information on non-geographical or socioeconomic characteristics of other localities may thus play an important role in selecting the reference jurisdictions. However, there are two difficulties. First, there are no *a priori* compelling reasons to choose specific socioeconomic factors that localities actually use when they identify their competitors.⁵ Second, faced with considerable amounts of information, local governments may be unable to find non-geographical neighbors that are relevant to them. Local governments may simply lose sight of their competitors amidst the plethora of shared information on the fiscal

⁵ Indeed, the choice over such factors is arbitrary in the empirical literature. Case et al. (1993) propose constructing \mathbf{W} with $W_{ij} = 1/|Q_i - Q_j|$, where Q is a relevant socio-economic variable. Typically, studies utilize the following variables for Q : (i) per capita income (Case et al. 1993, Boarnet and Glazer 2002, Finney and Yoon 2003, Baicker 2005, Caldeira 2012), (ii) population (Case et al. 1993, Rincke 2010), (iii) racial composition (Case et al. 1993, Boarnet and Glazer 2002), (iv) migration and commuting (Figlio et al. 1999, Baicker 2005, Rincke 2010), and (v) partisan affiliation (Foucault et al. 2008).

performance of numerous localities. It may then follow that a mere provision of fiscal information may not be good enough to foster yardstick competition.

The demarcation of “neighborhood” thus becomes crucial in the presence of costless information sharing. The central government could then help localities compete against each other not only by disseminating fiscal information among them for their performance comparison but also by demarcating their neighborhoods of competition, to help them find relevant benchmarks for comparison. As we argued in the Introduction, the FITS-M assume such dual roles in Japan.

The FITS-M categorize municipalities with the following criteria. The system of Japanese local administration classifies municipalities into cities, towns, and villages, and further categorizes cities into five types (designated cities, core cities, special cities, special wards in Tokyo metropolitan area, and ordinary cities).⁶ The FITS-M treat each of the first four types of cities as a single group. They also classify the other cities into 16 groups, and towns and villages into 15 groups, by their population n and the shares of local working population in the secondary sector (s_2) and the tertiary sector (s_3).⁷ The FITS-M thus categorize municipalities into 35 groups of “similar entities,” and provide a number of fiscal information indices on each group for performance comparison.⁸

⁶ The expenditure functions assigned to municipalities are identical except that towns and villages do not implement some of the social programs provided by cities and there are some variations among the five types of cities. Prefectures devolve parts of their expenditure functions to the first three types of cities, with the largest degree of devolution to designated cities, followed by core and then special cities. The functions assigned to the special wards are more or less similar to those of ordinary cities, although the Tokyo Metropolitan Government handles a few of the standard municipal functions (firefighting, water supply, and sewage disposal) for the special wards. Populations of designated cities are the largest among municipalities (3.7 to 0.71 million) followed, with some overlap of population ranges, by core cities (0.62 to 0.27 million) and special cities (0.57 to 0.19 million). Populations of the special wards in Tokyo metropolitan area range from 41.8 to 0.04 million.

⁷ There are four ranges of n for the other cities ($n < 50,000$; $50,000 \leq n < 100,000$; $100,000 \leq n < 150,000$; $150,000 \leq n$) and five for towns and villages ($n < 5,000$; $5,000 \leq n < 10,000$; $10,000 \leq n < 15,000$; $15,000 \leq n < 20,000$; $20,000 \leq n$). Meanwhile, the classification with s_2 and s_3 comprises four categories for the other cities ($s_2 + s_3 < 0.95$ and $s_3 < 0.55$; $s_2 + s_3 < 0.95$ and $0.55 \leq s_3$; $0.95 \leq s_2 + s_3$ and $s_3 < 0.65$; $0.95 \leq s_2 + s_3$ and $0.65 \leq s_3$) and three categories for towns and villages ($s_2 + s_3 < 0.80$; $0.80 \leq s_2 + s_3$ and $s_3 < 0.55$; $0.80 \leq s_2 + s_3$ and $0.55 \leq s_3$). The combination of these ranges and categories yields 16 (4×4) groups of ordinary cities and 15 (5×3) groups of towns and villages.

⁸ Such information includes within-group per capita averages and their annual changes of revenues by sources (e.g., taxes, transfers, and local bonds), expenses by type (e.g., personnel, personal transfers, and debt services), expenses by objective (e.g., social protection, public health, and public works), and capital expenses by funding sources (e.g., categorical grants, local bonds, and general revenues). The FITS-M also show the shares of these items, some of which can be used as indices for fiscal rigidity (the share of obligatory expenses), fiscal capacity, or self-sufficiency (the share of own revenues). Additional tables offer debt-related indices (e.g., several versions of debt ratio) and employment-related indices (e.g., wage level for municipal employees and per capita municipal employments).

We construct \mathbf{W} that characterizes fiscal interaction based on the FITS-M with the following un-normalized dichotomous element:

$$W_{ij} = 1 \begin{bmatrix} \text{Municipality } i \text{ belongs to an FIT-M group} \\ \text{that municipality } j \text{ belongs to} \end{bmatrix}. \quad (5)$$

Our purpose in the ensuing empirical exercise is to find how \mathbf{W} based on Eq. (5) fares against the standard \mathbf{W} 's based on geographical proximity. If the former performs better than the latter, we may argue that yardstick competition shapes the fiscal interaction in Japan. However, if municipalities within a given FITS-M category are geographically close for all FIT-M categories, a good performance of the model based on Eq. (5) may not necessarily imply the existence of yardstick competition. We thus examine if “neighbors” in a given FITS-M category are geographically close to each other, by plotting them on a map of Japan. Except those in the category of special wards in the Tokyo metropolitan area, we indeed find that these “neighbors” are located sparsely, with little geographical proximity, as shown in **Figure 1** with a selection of six categories of “similar localities.”

Figure 1

3. Empirical Implementation

3.1. Model

We utilize the following version of spatial autoregressive model:

$$g_{it} = \rho \cdot \left(\sum_{j=1}^N w_{ij} \cdot g_{jt} \right) + \sum_{k=1}^K \gamma_k \cdot x_{k,it} + \alpha_i + \tau_t + u_{it}, \quad (6)$$

where i and t respectively index the municipality and year; g_{it} is per capita municipal spending; $x_{k,it}$'s are K control variables; α_i is the municipal fixed effect for $i = 1, \dots, N$; τ_t denotes the year fixed effects for $t = 1, \dots, T$; and u_{it} is the error term. The fixed effects α_i allow for unobserved heterogeneity that affects municipal spending level, including the differences in expenditure function among the types of municipalities. As discussed in the previous section, w_{ij} is the $[i, j]$ -th element of an $N \times N$ spatial weights matrix \mathbf{W} with zero diagonal elements. We assume a spatial dependence in the error term, and specify it as another spatial auto-regressive model:

$$u_{it} = \lambda \cdot \sum_{i=1}^N m_{ij} \cdot u_{jt} + \varepsilon_{it}, \quad (7)$$

where m_{ij} is the $[i, j]$ -th element of another $N \times N$ spatial weights matrix \mathbf{M} , whose diagonal elements are also zero. We characterize the structure of \mathbf{M} by its un-normalized M_{ij} (defined analogously to W_{ij} in \mathbf{W}), which yields the row-normalized element $m_{ij} \equiv M_{ij}/\sum_j M_{ij}$. To distinguish between these two types of weights matrices, we shall call \mathbf{W} the interaction matrix and \mathbf{M} the disturbance matrix.

In spatial econometrics terminology, we see that Eq. (6) and Eq. (7) together yield a spatial autoregressive model with autoregressive disturbances (SARAR or SAC):

$$g_{it} = \rho \cdot \left(\sum_{j \neq i} w_{ij} \cdot g_{jt} \right) + \sum_k \gamma_k \cdot x_{k,it} + \alpha_i + \tau_t + \lambda \cdot \sum_{j \neq i} m_{ij} \cdot u_{jt} + \varepsilon_{it}.^9 \quad (8a)$$

We can express Eq. (8a) as a matrix expression:

$$\mathbf{g}_t = \rho \cdot \mathbf{W} \mathbf{g}_t + \mathbf{X}_t \boldsymbol{\gamma} + \boldsymbol{\alpha} + \boldsymbol{\tau}_t \cdot \mathbf{1} + \lambda \cdot \mathbf{M} \mathbf{u}_t + \boldsymbol{\varepsilon}_t, \quad (8b)$$

for $t = 1, \dots, T$, where $\mathbf{g}_t \equiv [g_{t1}, \dots, g_{tN}]'$, $\mathbf{X}_t \equiv [\mathbf{x}_{1t}, \dots, \mathbf{x}_{kt}, \dots, \mathbf{x}_{Nt}]$ with $\mathbf{x}_{kt} \equiv [x_{k,1t}, \dots, x_{k,Nt}]'$, $\boldsymbol{\alpha}_t \equiv [\alpha_1, \dots, \alpha_N]'$, $\mathbf{1} \equiv [1, \dots, 1]'$, $\mathbf{u}_t \equiv [u_{t1}, \dots, u_{tN}]'$, and $\boldsymbol{\varepsilon}_t \equiv [\varepsilon_{t1}, \dots, \varepsilon_{tN}]'$. We estimate Eq. (8) by the method of maximum likelihood (ML), basing our model on Lee and Yu (2010a, b) who assume that $\boldsymbol{\varepsilon}$ follows an i.i.d. normal distribution with constant standard deviation σ . We stack up Eq. (8b) over $t = 1, \dots, T$ and solve this system of equations to yield $N \times T$ reduced form equations for $\mathbf{g} \equiv \text{vec}[\mathbf{g}_1, \dots, \mathbf{g}_T]$ with accompanying multivariate errors. The ML function draws on these multivariate errors for a pair of \mathbf{W} and \mathbf{M} .¹⁰ We rely on the fixed- T asymptotic since our sample consists of large N and small T observations.

3.2. Sample and institutional backgrounds

We use a panel of Japanese municipalities as a sample for the estimation. Their fiscal

⁹ Unlike the spatial Durbin model, Eq. (8) excludes as regressors weighted values of other municipalities' control variables ($\sum_{j \neq i} w_{ij} x_{k,jt} \forall k$). This exclusion is due to the Nash assumption in the theoretical models of fiscal competition on which we base our arguments. The model assumes that the local government decides its fiscal variable g_i as an optimal response to a *given* value of \mathbf{g}_{-i} chosen by other governments and not to its controls ($x_{k,j} \forall k$ and $j \neq i$) that partially condition \mathbf{g}_{-i} .

¹⁰ We do not delineate the exact form of the ML function on which we perform optimization. Readers may easily refer to Yu et al. (2008) and Lee and Yu (2010a, b) to obtain appropriate guidance and explanation in this regard. We base our inference on what Lee and Yu call the "transformation approach." To actually obtain the estimates, we use XSMLE, a Stata module for spatial panel data model estimation introduced by Belotti et al. (2013). Since XSMLE produces the estimates based on the "direct approach," we adjust their values so that our inferences can be based on the transformation approach.

discretions along with the FITS-M system constitute a good case to examine intergovernmental interaction. Japan has two levels of local government: prefectures (upper level) and municipalities (lower level). Relevant national laws assign most of expenditure functions to local governments such that there is little overlap between the two levels. Municipalities are responsible for providing the variety of public services and social benefits, which include schooling (primary and lower secondary education), infrastructure (planning, construction, and maintenance), social protection (childcare, elderly care, and social assistance), fire protection, public health, garbage collection, water supply, and sewage disposal. While the central government sets national standards for a majority of these services and benefits, municipalities can plan their spending at their discretion in the sense that they can provide specific services and benefits beyond the base level set by the central government, and/or extend the scope of their beneficiaries.

Municipalities can also set their tax rates and fees/charges at the margin at their own discretion. Their finances in aggregate come from taxes (34.1%), general grants (16.9%), categorical grants (14.9%), municipal bond issuance (9.6%), and others items including fees/charges (24.5%).¹¹ A national law (Local Tax Law) assigns a specific set of taxes municipalities can collect, including taxes on land, housing and depreciable business assets (Fixed Property Taxes or “FTP”), and municipal taxes on individual and corporate incomes (Inhabitant Taxes or “IT”). While the law stipulates the standard rates, municipalities can set rates that differ from the standards. First, while there is a ceiling on the IT rate on corporate income, municipalities are free to increase the rates for IT on individual income and FPT beyond the relevant ceilings. In aggregate, these two taxes comprised 84% of total municipal tax revenues in FY2011. Second, municipalities can also reduce their tax rates below the standards despite the danger of paying fiscal penalties to the central government for such reductions.

Our sample consists of annual data on 1,637 municipalities¹² over three fiscal

¹¹ All figures are those for FY2010.

¹² At the end of FY2010, there were 1,750 municipalities. However, during FY2008 to FY2010, there were 96 instances of municipal mergers. We exclude municipalities that vanished on account of these mergers from the sample. Other than the four largest main islands (Hokkaido, Honshu, Shikoku, and Kyushu) and the Okinawa Islands, we also exclude 51 “island municipalities,” which have no geographical neighbors as they consist of only small islands. Furthermore, we exclude 10 municipalities that were hit by the Great Eastern Japan Earthquake in late FY2010 (i.e., March 2011) and 6 cities that changed their city classification type. These exclusions reduce the size of our sample to 1,637.

years (FY2008–FY2009). We choose this period on the following grounds. First, it is the period after the Japanese government completed decentralization measures on local taxes in the mid-2000s, when it performed tax-point transfers from national personal income tax to IT on personal income, and lifted ceilings on municipal tax rates for FPT and IT on personal income. Second, the central government made no changes to the criteria of the FIT-M grouping during this period. This constancy is important for maintaining the exogeneity of \mathbf{W} , as we will explain below.

3.3. Spatial weights matrices

Focusing on the effects of the interaction matrix \mathbf{W} based on the FITS-M (FM), we consider multiple types of spatial weights matrices. The most standard spatial weights matrices in the literature are the geographical proximity-based matrices, where the value of w_{ij} or m_{ij} decreases as distance d_{ij} between i and j increases. There are two major types of such matrices. One is the contiguity matrix (CG), and it formulates W_{ij} or M_{ij} as a dichotomous function that takes unity if i shares its border with j , and zero otherwise.¹³ The other is the inverse-distance matrix (ID) that treats W_{ij} or M_{ij} as a continuous function of d_{ij} , namely, W_{ij} (M_{ij}) = $1/d_{ij}$.¹⁴ If the spatial disturbance process occurs due to measurement errors for observations in contiguous spatial units,¹⁵ these distance-based matrices should fit the model better. In the following section, we examine how the FM fares against these distance-based alternatives, the CG and the ID.

The estimation of spatial regression models typically assumes that spatial weights are exogenous. While it is reasonable to assume so when we use the CG or the ID, it may be invalid to assume such exogeneity when the weights measure “socioeconomic” distance (Anselin and Bera 1998). For example, if we utilize variables like regional output or population to measure the distance as in Case et al. (1993), the weights would be endogenous since local public spending (the dependent variable) is likely to affect

¹³ A number of studies use the contiguity matrix as a baseline spatial matrix (Case et al. 1993, Boarnet and Glazer 2002, Hanes 2002, Revelli 2003, 2006, Geys 2006, Lundberg 2006, Revelli and Tovmo 2007, Werck et al. 2008, Nogare and Galizzi 2011, Bartolini and Santolini 2012, Caldeira 2012, Gebremariam et al. 2012, Costa et al. 2015).

¹⁴ The general form of the ID is the “distance decay” function specified as $W_{ij} = 1/d_{ij}^\delta$, where δ is some positive parameter (Murdock et al. 1993, Finney and Yoon 2003, Baicker 2005, Foucault et al. 2008, Caldeira 2012, Akai and Suhara 2013, Costa et al. 2015). Evidently, our ID assumes that $\delta = 1$. Variations of distance decay include distance decay with threshold D , where $W_{ij} = 1/d_{ij}^\delta$ if $d_{ij} < D$, and zero otherwise (Hanes 2002, Baicker 2005, Solé-Ollé 2006, Gebremariam et al. 2012, Costa et al. 2015).

¹⁵ See the classic discussion by Anselin (1988).

such variables. The literature has only recently started to offer methods that allow for the issues associated with endogenous weights (Kelejian and Piras 2014, Bhattacharjee et al. 2015, Qu and Lee 2015).

The FM depends on the FITS-M grouping, which, as we have seen, depends on types of municipalities, municipal population (n), and municipal industrial composition (s_2 and s_3). With respect to the latter two, we might plausibly suspect a reverse causation from the dependent variable to the FM, which violates the exogenous assumption. However, the data used for n , s_1 , and s_2 by the FITS-M are *constant* for our sample period (FY2008–FY2010). The FITS-M source the data for n , s_1 , and s_2 from the latest National Census, which is conducted every five years and, in those three fiscal years, it grouped municipalities using the values obtained from the census conducted in 2005. Thus, endogeneity may not be a serious issue for our weights matrices.

3.4. Control variables and data descriptions

Based on the standard theoretical model of local expenditure (e.g., Bergstrom and Goodman 1973), we include (i) regional income and (ii) central grants in the set of control variables. For the former, we use per capita income in the private sector aggregated at the municipal level. For the latter, we only consider general transfers and exclude other types of grants that are typically matching and categorical, since they affect local expenditure through changes in their matching *rates*, not their total *amounts*, which apply *uniformly* to all Japanese municipalities. When the uniform rates change over the years, year fixed effects, τ_t , could well capture their effects. Other controls include (iii) population, (iv) municipal surface area, (v) proportion of the population aged below 15, and (vi) proportion of the population aged 65 or older. The selection of these variables should be uncontroversial (e.g., Case et al. 1993). Population and surface area capture the possibility of potential congestion effects, including spatial costs and/or scale economies in the provision of municipal services. Localities with different age structures may have different demands for public services. The municipal fixed effects, α_i , account for any factors that differ across municipalities but remain constant over the years. Meanwhile, the year fixed effects, τ_t , allow for any unspecified annual changes that affect all municipalities equally in a given year. As mentioned, such changes include those in the matching elements of central grants.

Table 1 shows the descriptive statistics for the variables we use in our estimation.

We express all yen figures on a per capita basis (in 10 thousand yen). For the dependent variable, we use total municipal spending per municipal population. Among the control variables, we surrogate (i) regional income (per capita private sector income aggregated at the municipal level) with per capita taxable income aggregated at the municipal level, which is the tax base of municipal IT and the only income-related indicator available at the municipal level. For (ii) central grants, as mentioned above, we use only the amounts of general transfers (the Local Allocation Tax) and express it in per capita. The data for the other controls are straightforward.

We draw the data for municipal expenditures, central grants, municipal population, municipal surface areas, and populations of the two age groups from the *Annual Report on Municipal Finances* [*Shi-Cho-Son betsu Kessan Joukyo Shirabe*] for the relevant years (Ministry of Internal Affairs and Communications 2014a). We obtain municipal taxable income from the Regional Information Files in the Nikkei Economic Electronic Databank System (Nikkei Digital Media 2014).

Table 1

To construct the FM, we utilize the information from Ministry of Internal Affairs and Communication (2014b). For the weights based on geographical proximity (the CG and the ID), we utilize detailed geographical information on all municipalities. For the CG, we use the information obtained from the *National Database on Municipal Borders* [*Zenkoku Shi-Cho-Son Kai Data*], which is freely obtainable from ESRI Japan (2011). ESRI Japan constructs the database by processing information from the Database System on Local Jurisdiction Borders [*Kokudo Suchi Joho Gyosei Kuiki (Men) Deta*], maintained by the Ministry of Land, Infrastructure, Transportation and Tourism. Meanwhile, for the ID, we use the data on the latitudes and longitudes of municipal office locations (Geospatial Information Authority of Japan 2014) to calculate distance between a given pair of municipalities.

4. Results

This section estimates Eq. (8) with alternative specifications of \mathbf{W} and \mathbf{M} to assess if the information provided by the FITS-M shapes spending interaction among municipalities in Japan. **Table 2** shows the results for nine combinations of spatial weights matrices for

W and **M**. The first (FI, FC, and FF), second (II, IC, and IF), and the last (CI, CC, and CF) sets of three columns list the results with the FM, the ID, and the CG respectively for **W**. Within a set of three columns that share the same **W**, each column shows the results with the FM, the ID, and the CG respectively for **M**.

Table 2

4.1. Slope of the reaction function

Our key coefficient is ρ on the fiscal interaction term \mathbf{Wg} , which expresses the slope of the reaction function with which municipalities set their spending. For all nine models in **Table 2**, the estimates are positive and statistically significant, in line with expectations as per the literature. First, the theoretical predictions do not contradict the positive slope. We cannot *generally* predict its sign from the theoretical analysis of the benefit-spillover or resource-flow models.¹⁶ On the other hand, we can present a case of the positive slope from a theoretical model of yardstick competition. Besley and Case (1995) theoretically show that governments *mimic* one another, changing their policy *in the same direction* as their competitors. While their analysis only concerns tax interaction, we may readily extend it to spending interaction (Calderia 2012). Second, a majority of studies on expenditure competition estimate positive slopes.¹⁷ Some studies do find negative slopes. However, such cases approximate the benefit-spillover model, not the model of yardstick competition. Furthermore, they concern specific public services such as rescue (Hanes 2002), libraries (Finney and Yoon 2003), and cultural services (Lundberg 2006, Akai and Suhara 2013).

4.2. Comparing among different weights

Recall that our task is to investigate how models with the FM fare against those with the ID or CG. However, we cannot differentiate them by simply looking at their estimates

¹⁶ See Case et al. (1993, p. 298) for the benefit-spillover model and Brueckner (2003, pp. 180–181) for the resource-flow model. Using a specific form of benefit spillover, some studies associate the negative sign of ρ with free-riding behavior of local governments (Murdoch et al. 1993, Finney and Yoon 2003, Akai and Suhara 2013). However, in the presence of a general form of benefit spillover, we cannot generally determine the sign of the slope of the reaction function.

¹⁷ See Case et al. (1993), Bivand and Szymanski (2000), Boarnet and Glazer (2002), Revelli (2003), Baicker (2005), Dahlberg and Edmark (2008), Foucault et al. (2008), Werk et al. (2008), Ermini and Santolini (2010), Rincke (2010), Nogare and Galizzi (2011), Bartolini and Santolini (2012), and Costa et al. (2015). In addition, a number of empirical studies on tax competition provide analogous results. However, Chirinko and Wilson (2008/2013) show that, after controlling for spatio-temporal aggregate shocks and delayed responses, the reaction function indeed slopes *down* in tax competition.

for ρ , as their values are all positive and statistically significant for the nine models, implying the existence of fiscal interaction without violating the theoretical prediction. Nonetheless, the statistical significance does not necessarily imply that all the models are equally satisfactory. It only allows us to test the hypothesis that $\rho = 0$ for each model, from which we cannot compare their performance. Furthermore, if we misspecify the model with a wrong choice of \mathbf{W} , the estimate for ρ may converge to some nonzero value that is different from its true value. Indeed, as MacKinnon (1983) argues, a regression model that at first glance seemed to be satisfactory might turn out to be misleading on closer investigation. We therefore examine the models in **Table 2** more closely by performing a robustness check and fictive non-nested hypothesis tests, and by comparing the information criteria and likelihood dominance criterion.

4.2.1 Robustness check on coefficient estimate for fiscal interaction

A robustness check examines how a coefficient estimate in question behaves as we modify the regression specification in some way (Lu and White 2014). If the estimate is “fragile” (i.e., sensitive to the modifications), we may suspect that the model is misspecified (Leamer 1983). Our robustness check examines how the estimate for ρ based on a given choice of \mathbf{W} changes over different choices of \mathbf{M} , which should typically be relevant in our context. The literature often argues that if we estimate the model without allowing for the spatial error dependence that is in fact present, the result may give a false impression of fiscal interaction when none is in fact occurring (Case et al. 1993, Brueckner 1998, Brueckner and Saavedra 2000, Revelli 2001). Extending this line of reasoning, we could argue that when we correctly specify the fiscal interaction with a correct \mathbf{W} , the ρ estimate would not change much over different choices of spatial error dependence (i.e., different \mathbf{M} s). On the other hand, when we misspecify the model with a wrong \mathbf{W} , it would change drastically over different \mathbf{M} s.

As seen from **Table 2**, the robustness check encourages us to choose the FM for \mathbf{W} . For the models with the FM for \mathbf{W} (FI, FC, and FF), the ρ estimates are quite stable against the three choices of \mathbf{M} , differing only up to the third decimal place (.439, .437, and .431). In contrast, the estimates for the models with the CG for \mathbf{W} (CI, CC, and CF) are fragile in the sense that its estimates change drastically (.081, .457, and .093). The estimates for the models with the ID for \mathbf{W} (II, IC, and IF) are not as fragile as those with the CG (.595, .638, and .581), but they are also not as stable as those with the FM.

It may therefore be safe to maintain that the FM survive the robustness check for \mathbf{W} .

4.2.2 Fictive non-nested hypothesis tests

Note that we cannot nest one of the nine models, say, $g = h(\boldsymbol{\phi}) + \varepsilon$ within another model $g = f(\boldsymbol{\gamma}) + \varepsilon$, with $\boldsymbol{\phi}$ and $\boldsymbol{\gamma}$ being their respective parameters. We therefore have to utilize non-nested hypothesis testing (e.g., MacKinnon 1983, 1992; Pesaran and Weeks 2003) if we are to evaluate the nine models in **Table 2**. Typically, non-nested hypothesis testing starts by artificially nesting two competing models into a composite model:

$$g_{it} = (1 - \theta) \cdot f(\boldsymbol{\phi}) + \theta \cdot h(\boldsymbol{\gamma}) + \varepsilon_{it}. \quad (9)$$

With this composite model, we can express either of the two models as its restricted form. For example, we may test $g = f(\boldsymbol{\phi}) + \varepsilon$ by imposing the restriction $\theta = 0$ on Eq. (9). Unfortunately, however, we cannot estimate Eq. (9), since we cannot separately identify θ , $\boldsymbol{\phi}$, and $\boldsymbol{\gamma}$ in Eq. (9).¹⁸ One of the standard methods to overcome this difficulty is to employ the *J*-test, which replaces parameters $\boldsymbol{\gamma}$ for the model that is not being tested with their consistent estimates when $g = h(\boldsymbol{\gamma}) + \varepsilon$ is true (cf., Davidson and MacKinnon 1981). Anselin (1984, 1986) was among the first who has taken advantage of the *J*-test to evaluate models with different forms of \mathbf{W} . Recent studies on spatial econometrics extend the *J*-test to be amenable to examining a variety of spatial models (Kelejian 2008, Burridge and Fingleton 2010; Kelejian and Piras 2011, 2015, Burridge 2012).

However, we may be able to reject a subset of the nine models to obtain a smaller set of viable models without actually estimating Eq. (9). We could do so through the following *fictive* non-nested hypothesis testing that only compares the individual log likelihoods of the nine models.¹⁹ When we artificially nest a pair of models in **Table 2** into a composite model in the form of Eq. (9), its log likelihood L always takes on a value that is larger than the larger of the log likelihood values of the two models. If we express the log likelihood of Model i as L_i , we find that L_{FI} takes on the largest value among the nine models, and of course, that $L > L_{\text{FI}}$. Then, we consider a quantity

$$LR_{\text{FI},j} = 2 \cdot (L_{\text{FI}} - L_j)$$

where j indexes each of the eight models other than Model FI. If $LR_{\text{FI},j} > \chi^2(13)$ for a

¹⁸ Note that both $h(\boldsymbol{\phi})$ and $h(\boldsymbol{\gamma})$ are *not* linear in parameters, as seen from Eq. (8a). They contain the non-white noise part of the spherical error u in Eq. (6), that is, $\lambda \sum_j m_j \mu_j$ of Eq. (7).

¹⁹ The following argument is analogous to the issue of choosing between logarithm and level forms of a dependent variable without estimating the Box–Cox form that nests the former two forms with additional parameters. For a textbook explanation, see Davidson and MacKinnon (1993, pp. 491–492).

given level of statistical significance, the likelihood-ratio test rejects Model j at that significance level, since its test statistics is $2 \cdot (L - L_j)$ and $L > L_{FI}$. From the table, we find that $LR_{FI, FC} = 10$, $LR_{FI, FF} = 18$, $LR_{FI, II} = 102$, $LR_{FI, IC} = 106$, $LR_{FI, IF} = 32$, $LR_{FI, CI} = 106$, $LR_{FI, CC} = 68$, and $LR_{FI, CC} = 34$. At the .05 level,²⁰ we can reject all the six models that do not use the FM for \mathbf{W} , whereas we cannot reject either Model FF or Model FC. The rejection may imply the shortcomings of the models being tested (Pesaran and Weeks 2003) and constitute evidence for their misspecification (MacKinnon 1983, 1992). With these estimates, therefore, we could and did narrow a set of surviving models down to Models FI, FC, and FF, all using the FM for \mathbf{W} .

4.2.3 Information criteria and likelihood dominance criterion

In the non-nested hypothesis testing, we may end up rejecting or accepting all the models under consideration, thus either choosing none of them or retaining multiple (all) models. If we could obtain the value of L , we might reject both Models FI and FC. However, this might not matter much if we frame our exercise as that of model selection, where we are supposed to find one model as the “best” one among the alternatives. For model selection, Granger et al. (1995) recommend using the information criteria, including Akaike’s information criteria (AIC) and the Bayesian information criteria (BIC). Not surprisingly, therefore, applied studies in spatial econometrics indeed utilize the AIC to evaluate multiple models with alternative spatial weights (Leenders 2002, Getis and Aldstadt 2004, Stakhovych and Bijmolt 2008).

In addition to these criteria, Pollak and Wales (1991) propose what they call the likelihood dominance criterion (LDC), which extends the above-mentioned fictive non-nested hypothesis testing. They show that, ruling out the possibilities of either rejecting or accepting *both* models, we can identify the accepted and rejected models by looking at their values of log likelihood and their numbers of parameters.²¹ In particular, if the two models have the same number of parameters, the non-nested test always accepts the model with the higher value of log likelihood and rejects the model with the smaller value. Thus, the ordering by the LDC always prefers the model with the highest log

²⁰ Since $q = 11$ in our case, the critical values of $\chi^2(11)$ are 17.3, 19.7, and 24.7 for the .10, .05, and .01 levels of significance, respectively.

²¹ In this respect, when models to be evaluated are not nested, we can interpret a model selection with the likelihood dominance criterion (LDC) as a ranking of multiple hypotheses through non-nested hypothesis testing that rules out the possibilities of either rejecting or accepting *all* models.

likelihood if the number of parameters is identical among the alternatives. A Monte Carlo study also shows that the LDC method outperforms non-nested hypothesis testing procedures in selecting true models (Saha et al. 1994).

We therefore use the AIC, BIC, and LDC (the log likelihood) to compare the nine models in **Table 2**. These three criteria yield the same ordering, favoring the three models with the FM for **W**. Among them, Model FI has the smallest values of the AIC and BIC, and the largest value of log likelihood. We thus select Model FI, which uses the FM for **W** and the ID for **M**.

4.3. Further analysis I: The effects of distance within a given FITS-M group

All these diagnostics lead to the conclusion that Model FI is the best model among the nine models in **Table 2**, corresponding to the FM for interaction weights and the ID for disturbance weights. This suggests that the FIT-M may play a primary role in facilitating yardstick competition through demarcating the boundary of neighbors and providing their fiscal information. The result is also consistent with the view that measurement errors over contiguous units motivate spatial error dependence, since we find the weights based on geographical proximity, ID, more appropriate for the error weights.

Even if the FITS-M play the primal role in fiscal competition, however, geography may still play a (secondary) role in differentiating the effects within a given FITS-M group. In particular, it may be natural to suspect that the more closely located a pair of municipalities in a given FITS-M group, the more closely they may keep an eye on each other. It should then be worthwhile to examine a hybrid spatial weights matrix for **W** that involves the indicators for both FITS-M grouping and geographical proximity. In general, when a hybrid weights matrix adjusts binary weights with a non-binary measure, the binary factor relates to the primary effect, while the adjusting measure relates to the secondary effect, since the latter effect applies only to entities that have a non-zero value in the binary factor.²² We could then capture the secondary effect of geographical proximity between a pair of municipalities in the same FITS-M group, by adjusting the

²² Several studies in the literature have indeed employed hybrid weights matrices. In particular, they use the matrices to differentiate the effects of interaction between a pair of localities whose geographical proximity is identical (contiguous). For example, Rincke (2010) adjusts the contiguity matrix with an index of commuting patterns between a pair of localities. Other studies adjust the contiguity index with some forms of population-related index (Werck et al. 2008, Nogare and Galizzi 2011, Akai and Suhara 2013, Costa et al. 2015). These weights capture the effect of the adjusting factor (commuting patterns or population characteristics) *after* allowing for the influence of geographical proximity (contiguity).

binary elements in the FM matrix with a measure for geographical proximity. We thus construct hybrid weights by multiplying the un-normalized elements in Eq. (5) with a proximity measure ϕ_{ij} between i and j :

$$W_{ij} = 1 \left[\begin{array}{l} \text{Municipality } i \text{ belongs to an FIT-M group} \\ \text{that municipality } j \text{ belongs to} \end{array} \right] \times \phi_{ij}. \quad (10)$$

Candidates for ϕ_{ij} are the inverse distance index and the contiguity index. In what follows, nonetheless, we use only the inverse distance index, that is, $\phi_{ij} = 1/d_{ij}$, since there are very few municipalities that are adjacent within a given FITS-M group (recall the panels in **Figure 1**), except the group of special wards in Tokyo Metropolitan area.

Table 3 lists the estimation results for the models with this hybrid weights matrix (the HB) for \mathbf{W} over different choices of \mathbf{M} (HI, HC, HF, and HH). For the sake of completeness, the table also includes the models with the HB for \mathbf{M} and each of the ID, the CG, and the FM for \mathbf{W} (IH, CH, and FH), which are not listed in **Table 2**. We thus have 16 models in total, nine models in **Table 2** and seven models in **Table 3**.

Table 3

We then analogously conduct the diagnostics to find the “best” model among these 16 models. First, the robustness check on the fiscal interaction coefficients shows that the HB may be as good a candidate as the FM for \mathbf{W} . The results for the models with the HB for \mathbf{W} (HI, HC, HF, and HH) indicate that their ρ estimates, with the standard deviation of .029 and the max-min difference of .058, are reasonably stable against the four different choices of \mathbf{M} (.389, .402, .376, and .434). Meanwhile, the models with the FM for \mathbf{W} now extend from the three models in **Table 2** to include Model FH in **Table 3** with the ρ estimate of .367. This addition yields the standard deviation of .022 and the max-min difference of .070 among the four values of ρ from Models FI, FC, FF, and FH. While the standard deviation is smaller for the four cases with the FM for \mathbf{W} , the max-min difference is smaller for the four cases with the HB for \mathbf{W} .

Second, the fictive non-nested hypothesis testing points to the four models with the HB for \mathbf{W} (HI, HC, HF, and HH) and the model with the FM for \mathbf{W} and the HB for \mathbf{M} (FH). They all have almost identical values of log likelihood (−11,217 for HI, −11,218 for HC, HF, and HH, and −11,219 for FH), which does not allow us to effectively differentiate the four models. Nonetheless, Model HI has the largest value of log

likelihood among the 16 models, which is large enough to falsify the other models, except Model FH, that do not use the HB for \mathbf{W} through the fictive non-nested hypothesis tests. Note that the exception, Model FH, has the FM for \mathbf{W} and the HB for \mathbf{M} , indicating the FITS-M still play an important role in fiscal interaction.

Third, we compare the 16 models with the AIC, BIC, and LDC. Both the AIC and BIC select Model HI. While the values of the log likelihood are almost identical among the five models that have the five largest values, the LDC also selects Model HI.

While the previous subsection suggests Model FI as the best model among the nine in **Table 2**, the additional diagnostics seem to point to Model HI as the best among the 16 models. The model has the hybrid spatial weights matrix (the HB) for the fiscal interaction and the proximity-based matrix (the ID) for the error dependence. This then implies that, while the FITS-M serve as the primary vehicle in facilitating yardstick competition among municipalities, localities located more closely exert more influence than the others within a given FITS-M group. In addition, this result may again lend support to the view that motivates the error dependence as measurement errors for observations in contiguous spatial units.

4.4. Further analysis II: Regionally differentiated temporal effects

As we have mentioned, the empirical studies on fiscal interaction often argue that if we estimate the model *without* allowing for the spatial error dependence that is in fact present, we may obtain a false impression of fiscal interaction when none is in fact occurring. The majority of the studies in the literature allow for the error dependence by specifying the dependence as Eq. (7), that is, a spatial autoregressive process with a single autoregressive parameter (λ) and a spatial weights matrix \mathbf{M} . Baily et al. (2015) call this specification “spatial model” and frame it as a subclass of (a more general) cross-section error dependence (CSD).

Another subclass of the CSD is what Baily et al. (2015) call “factor model.”²³ The model characterizes the CSD in terms of unobserved common temporal factors $\mathbf{f}_t = [f_t, \dots, f_t]'$ whose marginal effects or factor loadings $\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_N]'$ differ across the cross-section units. The total effect of the factors is then given as an additive term $\boldsymbol{\gamma} \mathbf{f}_t$ in a regression model, taking on different values over the cross-section units i and time

²³ The spatial model and the factor model are not exclusive. While we could substitute the spatial model with the factor model (e.g., Chirinko and Wilson 2008/2013), we could also employ it along with the spatial model (e.g., Holly et al. 2010).

periods t . When applied to our analysis, the factor model captures cases where common temporal shocks (common factors) affect all municipalities with differentiated marginal impacts (factor loadings) across them.²⁴

To allow for the factor model, Pesaran (2006) proposes the common correlated effects (CCE) estimator. One way to obtain the CCE estimates is to augment the regression model with additional regressors that consist of the interactions of cross-sectional dummies and the cross-sectional averages of the dependent and independent variables. However, we could confidently obtain the CCE estimates only when T is sufficiently large (Bailey et al. 2015). Since our sample is very short (with $T = 3$) and large (with $N = 1,637$),²⁵ we might have to content ourselves *only* with the use of the spatial model.²⁶ However, if we could assume that factor loadings take on a common value for a group of multiple adjacent municipalities within a wider area, we could allow for this type of unobserved temporal shock by augmenting the regression with the interactions of the time dummies and the dummies for those wider areas. Although this relaxed assumption is not a perfect substitute for the factor model, it may still be a viable alternative to control the temporal shocks if the municipalities are too small a unit to differentiate factor loadings.

We thus reestimate all 16 models in **Tables 2** and **3**, augmenting them with the interactions of the time and wider area dummies (“region–time dummies” for short). Our estimation thus allows for both the regionally differentiated temporal shocks (a substitute for the factor model) and the autoregressive process of the error term (the spatial model). We consider two choices of wider areas in Japan, 47 prefectures and 8 regional blocks. We thus obtain $47 \times 3 = 141$ and $8 \times 3 = 24$ interactions as the additional

²⁴ Obviously, the standard time effect [τ_t in Eq. (6)] cannot allow for these temporal effects. We also thought of including the interactions of time and cross-section dummies in the model but this was infeasible since the number of such interactions amounts to the sample size ($N \times T$).

²⁵ While K refers to the number of parameters to be estimated in the original (linear) model, the augmented model has $K + N \times (K + 1)$ parameters to be estimated. Given our sample with $N = 1,637$ and $T = 3$, this method is simply infeasible since $K + N \times (K + 1) > N \times T$. Chirinko and Wilson (2008/2013) also suggest a way to reduce the number of the augmented regressors by restricting parameters in the augmented model. However, this still requires $K + N$ parameters to be estimated. Since we only have $T = 3$, this may still be too large a number of parameters. Furthermore, as Chirinko and Wilson (2008/2013) report, this restriction necessitates a nonlinear estimation, which may have difficulty converging. Indeed, we did have difficulty in converging with the models that have as augmented regressors the interactions of prefectural and time dummies.

²⁶ Bailey et al. (2015) note, “Almost all spatial econometric models estimated in the literature assume that the spatial parameters do not vary across the units. ... Such parameter homogeneity is not avoidable when T is very small, but need not be imposed in the case of large panels where T is sufficiently large.”

regressors respectively for prefectures and regional blocks, barring those excluded because of collinearity. However, we face computational difficulties when using prefectural dummies, failing to obtain convergence in the parameter estimates. We thus only list the results for the cases with the less demanding choice of eight regional blocks in **Table 4**.

Table 4

Table 4 shows that, even after controlling the region–time dummies, the diagnostics select models with the HB for **W**, which shows that the FITS-M still play an important role in fiscal interaction. First, the robustness check on the ρ coefficients shows that the HB may be as good a candidate as the FM for **W**. Second, the fictive non-nested hypothesis testing rejects the eight models with either the ID or the CG for **W**,²⁷ pointing to the four models with the HB for **W** (HI', HC', HF', and HH') and the four models with the FM for **W** (FI', FC', FF', and FH'). Third, while the values of log likelihood are almost identical among the five models with the largest values, the AIC, BIC, and LDC all select Model HH', a model with the HB for both **W** and **M**. Therefore, now, the “best” choice of the spatial weights matrices is the combination of the HB for both **W** and **M** with the region–time dummies, while previously it was the combination of the HB for **W** and the ID for **M** in **Tables 2** and **3** without the dummies. However, this difference may not matter much. Since the spatial auto-regressive error is now statistically insignificant for all the four models that use the HB for **W**, the choice of **M** might be irrelevant in selecting among them. Perhaps, the inclusion of the dummies may be sufficient to allow for the spatial errors specified as Eq. (7).

4.5 Discussing the coefficients

We conclude our analysis by discussing the coefficient estimates in Model HI in **Table 3**, comparing them to the analogous estimates (Model HI') in **Table 4**. The coefficient for fiscal interaction is statistically significant and positive ($\beta = .389$). *Ceteris paribus*, a 100-yen increase in municipal spending by i 's neighbors in its FITS-M category increases i 's spending by 39 yen. Including the region-time dummies does not change the estimate ($\beta = .389$). Meanwhile, the coefficient for the disturbance interaction is also significant and positive ($\rho = .300$). As mentioned, the inclusion of the dummies makes

²⁷ Given the interactions between time and regional dummies, we now have $q = 25$. The critical values of $\chi^2(25)$ are then 34.4, 37.7, and 44.3 for the .10, .05, and .01 levels of significance, respectively.

the spatial correlation small ($\rho = -.158$) and insignificant.

Increases in income, central grants, and population induce municipal spending to expand, as we usually expect. On the other hand, an increase in surface area reduces spending. This might be due to the particular properties of the surface area data. Since our sample excludes municipalities that merged during the sample period, changes in jurisdictional area should be due to some erratic events whose reasons are difficult to identify. Note that, while the effects of the other three variables remain almost identical even with the inclusion of the region-time dummies, the inclusion makes the effect of income smaller and insignificant. We conjecture that there may exist a large correlation among private incomes (aggregated at the municipal level) within a regional block, which are subject to temporal common shocks at the regional level. This may indeed sound plausible, since local economies are not closed within municipal boundaries, but constitute a larger regional economy at the block level.

Meanwhile, the two demographic variables, namely, the proportions of the young and the elder, are not statistically significant. This may be because we use a very short panel (three years) of municipalities with municipal and year fixed effects. These variables may have changed little during the three years so that the municipal fixed effects absorb their entire effects. Even if some of them did change at all over the years, their variations may have been uniform across the municipalities so that the year effects presumably absorb such changes. As expected, the inclusion of the region-time dummies does not change the results.

We may elaborate somewhat on the effects of income and central grants. The coefficient on income is statistically positive but small (.069) without the inclusion of the region-time dummies, which is consistent with the values provided by surveys of empirical studies on local government expenditure (Hines and Thaler 1995, Bailey and Connolly 1998). The inclusion of the dummies makes its effect statistically insignificant, possibly for the reasons we have mentioned above. On the other hand, the coefficient on central grants is 1.833, which is quite large compared to the range of values found in the literature (.6–1.0) (Hines and Thaler 1995, Bailey and Connolly 1998). Our estimate implies that, *ceteris paribus*, municipalities spend roughly 1.8 yen for each yen obtained from the central government. Holding other variables constant, this value suggests that an increase in the municipal spending exceeds an increase in the grant it receives. This may due to the institutional characteristics in Japanese systems of local public finance

and intergovernmental fiscal relations, which are quite different from those of their North American or European counterparts. In addition, there may be methodological issues to be considered when estimating the effects of central grants.²⁸ Exploring this topic further is beyond the scope of this paper, however. We plan to study this important topic in our future research.

5. Concluding Remarks

The Japanese government provides local governments with information on local fiscal performance and categorizes them into groups of “similar localities” through the FITS-M. In doing so, it intends to let local governments use the information provided through the FITS-M as a benchmark for their fiscal planning. In other words, the FITS-M allow municipalities to find other municipalities in their “neighborhood” with which they fiscally compete and to use the fiscal information on these group members as a “yardstick.” In this paper, we took advantage of this unique system to estimate municipal spending function in Japan. In particular, we examined if the FITS-M helps identify a defensible spatial weights matrix that properly describes spending interaction among municipalities. We estimated 32 models that have different combinations of interaction and error weights and compared them to select the most appropriate model. Our analysis chose the models that use a hybrid weights matrix, which allows for both the effects of the FITS-M grouping and proximity within the group. It also showed that the models with the FITS-M matrix, which excludes the effect of proximity, also perform better than those with the traditional proximity-based weights matrices (contiguity and inverse distance). This result is consistent with the claim that the FITS-M work as intended. It also implies that geographical proximity only affects competition between a pair of municipalities within the same FITS-M group, but not competition between a pair of those that belong to different FITS-M groups even if they are located close by.

Based on our arguments in this paper, this result would then lend support to the view that spending interaction among Japanese municipalities originates from yardstick competition and not from other types of competitions based on resource flow or benefit spillover. Since this paper focused on total municipal spending, however, this statement

²⁸ A recent study by Leduc and Wilson (2015) nicely summarizes recent evidence on the flypaper effect and discusses methodological issues in estimating the effect of central grants.

may be relevant only to the description of fiscal interaction in terms of “average” spending. As briefly mentioned in Section 3.2, municipalities in Japan take on a number of different types of spending, some of which may cause forms of interaction that differ from yardstick competition. While our result would nonetheless imply that yardstick competition dominates other forms of interaction on average, it is indeed interesting to apply our procedure to a subcategory of municipal expense and examine how, within the subcategory, the FITS-M based weights matrix would fare against those based on geographical proximity. However, pursuing this line of examination is beyond the scope of this paper. We plan to tackle this important topic in our future research.

Acknowledgements

We are grateful to Robert Chirinko, Editor-in-Chief, and two anonymous reviewers for their helpful observations and constructive suggestions, which substantively improved the paper. We also thank Mutsumi Matsumoto for his valuable comments. Hayashi acknowledges financial support from the Japan Finance Organization for Municipalities and the Grant-in-Aid for Scientific Research (15H03359) of the Japan Society for the Promotion of Science.

References

- Ahrens, A., Bhattacharjee, A. (2015). Two-step Lasso estimation of the spatial weights matrix. *Econometrics*, 3(1), 128–155.
- Akai, N., Sahara, M. (2013). Strategic interaction among local governments in Japan: An application to cultural expenditure. *Japanese Economic Review*, 64(2), 232–247.
- Anselin, L. (1984). Specification tests on the structure of interaction in spatial econometric models. *Papers of the Regional Science Association*, 54(1), 165–182.
- Anselin, L. (1986). Non-nested tests on the weight structure in spatial autoregressive models: Some Monte-Carlo results. *Journal of Regional Science*, 26(2), 267–284.
- Anselin, L. (1988). *Spatial econometrics: Methods and models*. Dordrecht: Kluwer Academic Publishers.
- Anselin, L., Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah, D. E. Giles (Eds.), *Handbook of Applied Economic Statistics* (pp. 237–290), New York: Marcel Dekker.
- Baicker, K. (2005). The spillover effects of state spending. *Journal of Public Economics*, 89(2/3), 529–544.
- Bailey, S. J., Connolly, S. (1998). The flypaper effect: Identifying areas for further research. *Public Choice*, 95(3/4), 335–361.
- Bailey, N., Holly, S., Pesaran, M. H. (2015). A two-stage approach to spatiotemporal analysis with

- strong and weak cross-sectional dependence. *Journal of Applied Econometrics*, doi: 10.1002/jae.2468.
- Bartolini, D., Santolini, R. (2012). Political yardstick competition among Italian municipalities on spending decisions. *Annals of Regional Science*, 49(1), 213–235.
- Beenstock, M., Felsenstein, D. (2012). Nonparametric estimation of the spatial connectivity matrix using spatial panel data. *Geographical Analysis*, 44(4), 386–397.
- Belotti, F., Hughes, G., Mortari, A. P. (2013). XSMLE: Stata module for spatial panel data models estimation. Statistical Software Components S457610, Boston College Department of Economics, Revised March 15, 2014.
- Bergstrom, T. C., Goodman, R. P. (1973). Private demands for public goods. *American Economic Review*, 63(3), 280–296.
- Besley, T., Case, A. (1995). Incumbent behavior: Vote seeking, tax setting and yardstick competition. *American Economic Review*, 85(1), 25–45.
- Bhattacharjee, A., Holly, S. (2013). Understanding interactions in social networks and committees. *Spatial Economic Analysis*, 8(1), 23–53,
- Bhattacharjee, A., Jensen-Butler, C. (2013). Estimation of the spatial weights matrix under structural constraints. *Regional Science and Urban Economics*, 43(4) 617–634.
- Bhattacharjee, A., Castro, E., Maiti, T., Marques, J. (2015). Endogenous spatial structure and delineation of submarkets: A new framework with application to housing markets. *Journal of Applied Econometrics*, doi: 10.1002/jae.2478.
- Bivand, R., Szymanski, S. (1997). Spatial dependence through local yardstick competition: Theory and testing. *Economics Letters*, 55(2), 257–265.
- Bivand, R., Szymanski, S. (2000). Modeling the spatial impact of the introduction of Compulsory Competitive Tendering. *Regional Science and Urban Economics*, 30(2), 203–219
- Boadway, R. (1983). On the method of taxation and the provision of local public goods: Comment. *American Economic Review*, 72(4), 846–851.
- Boarnet, M., Glazer, A. (2002). Federal grants and yardstick competition. *Journal of Urban Economics*, 52(1), 53–64.
- Boarnet, M., Cemiglia, F., Revelli, F. (2004). Yardstick competition in intergovernmental relationships: Theory and empirical predictions. *Economics Letters*, 83(3), 325–333.
- Boots, B., Dufournaud, C. (1994). A programming approach to minimizing and maximizing spatial autocorrelation statistics. *Geographical Analysis*, 26(1), 54–66.
- Brueckner, J. K. (1988). Testing for strategic interaction among local governments: The case of growth controls. *Journal of Urban Economics*, 44(3), 438–467.
- Brueckner, J. K. (2003). Strategic interaction among governments: An overview of empirical studies. *International Regional Science Review*, 26(2), 175–188.
- Brueckner, J. K., Saavedra, L. A. (2001). Do local governments engage in strategic property-tax competition? *National Tax Journal*, 54(3), 203–230.
- Burridge, P. (2012). Improving the J test in the SARAR model by likelihood-based estimation. *Spatial Economic Analysis*, 7(1), 75–107.
- Burridge, P., Fingleton, B. (2010). Bootstrap inference in spatial econometrics: the J-test. *Spatial Economic Analysis*, 5(1), 93–119.
- Caldeira, E. (2012). Yardstick competition in a federation: Theory and evidence from China. *China Economic Review*, 23(4), 878–897.
- Case, A. C., Rosen, H. S., Hines Jr., J. R. (1993). Budget spillovers and fiscal policy interdependence. *Journal of Public Economics*, 52(3), 285–307.
- Chirinko, R. S., Wilson, D. J. (2008/2013). Tax competition among U.S. states: Racing to the bottom or riding on a seesaw? Federal Reserve Bank of San Francisco Working Paper Series 2008-03.
- Costa, H., Veiga, L. G., Portela, M. (2015). Interactions in local governments' spending decisions:

- Evidence from Portugal. *Regional Studies*, 49(9), 1441–1456.
- Dahlberg, M., Edmark, K. (2008). Is there a “race-to-the-bottom” in the setting of welfare benefit levels? Evidence from a policy intervention. *Journal of Public Economics*, 92(5), 1193–1209.
- Davidson, R., MacKinnon, J. G. (1981). Several tests for model specification in the presence of alternative hypotheses. *Econometrica*, 49(3), 781–793.
- Ermini, B., Santolini, R. (2010). Local expenditure interaction in Italian municipalities: Do local council partnerships make a difference? *Local Government Studies*, 36(5), 655–677.
- ESRI Japan (2011). The national database on municipal borders [Zenkoku shi-cho-son kai data], Version 7.1. <http://www.esrij.com/products/japan-shp/>. Accessed 12 November 2014.
- Facchini, F. (2014). The determinants of public spending: a survey in a methodological perspective. MPRA Paper No. 53006, Munich Personal RePEc Archive.
- Figlio, D. N., Kolpin, V. W., Reid, W. E. (1999). Do states play welfare games? *Journal of Urban Economics*, 46(3), 437–454.
- Finney, M. M., Yoon, M. J. (2003). Asymmetric interdependence in the provision of a local public good: An empirical examination. *Public Finance Review*, 31(6), 648–668.
- Folmer, H., Oud, J. (2008). How to get rid of W: A latent variables approach to modelling spatially lagged variables. *Environment and Planning A*, 40(10), 2526–2538.
- Foucault, M., Madies, T., Paty, S. (2008). Public spending interactions and local politics: Empirical evidence from French municipalities. *Public Choice*, 137(1), 57–80.
- Gebremariam, G. H., Gebremariam, T. G., Schaeffer, P. V. (2012). County-level determinants of local public services in Appalachia: A multivariate spatial autoregressive model approach. *Annals of Regional Science*, 49(1), 175–190.
- Geospatial Information Authority of Japan, 2014. Latitude and longitude of prefectural and municipal office locations [To-do-fu-ken Shi-cho-son no To-zai-nan-boku Ten no Keido Ido]. <http://www.gsi.go.jp/KOKUJYOHO/center.htm>. Accessed 12 November 2014.
- Getis, A., Aldstadt, J. (2004). Constructing the spatial weights matrix using a local statistic. *Geographical Analysis*, 36(2), 90–104.
- Geys, B. (2006). Looking across borders: A test of spatial policy interdependence using local government efficiency ratings. *Journal of Urban Economics*, 60(3), 443–462.
- Granger, C. W. J., King, M. L., White, H. (1995). Comments on testing economic theories and the use of model selection criteria. *Journal of Econometrics*, 67(1), 173–187.
- Griffith, A. D., Lagona, R. (1998). On the quality of likelihood-based estimators in spatial autoregressive models when the data dependence structure is misspecified. *Journal of Statistical Planning and Inference*, 69(1), 153–174.
- Hanes, N. (2002). Spatial spillover effects in the Swedish local rescue services. *Regional Studies*, 36(5), 531–539.
- Harris, R., Moffat, J., Kravtsova, V. (2011). In search of “W.” *Spatial Economic Analysis*, 6(3), 249–270.
- Hines, J. R., Thaler, R. H. (1995). Anomalies: The flypaper effect. *Journal of Economic Perspectives*, 9(4), 217–226.
- Holly, S., Pesaran, M. H., Yamagata, T. (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics*, 158(1), 160–173.
- Kelejjan, H.H. (2008). A spatial J-test for model specification against a single or set of non-nested alternatives. *Letters in Spatial and Resource Sciences*, 1(1), 3–11.
- Kelejjan, H. H., Piras, G. (2011). An extension of Kelejjan’s J-test for non-nested spatial models. *Regional Science and Urban Economics*, 41(3), 281–292.
- Kelejjan, H. H., Piras, G. (2014). Estimation of spatial models with endogenous weighting matrices and an application to a demand model for cigarettes. *Regional Science and Urban Economics*, 46(C), 140–149.

- Kelejjan, H. H., Piras, G. (2015). An extension of the J-test to a spatial panel data framework. *Journal of Applied Econometrics*, doi: 10.1002/jae.2425.
- Leamer, E. E. (1983). Let's take the con out of econometrics. *American Economic Review*, 73(1), 31–43.
- Lee, L. F., Yu, J. (2010a). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154(2), 165–185.
- Lee, L. F., Yu, J. (2010b). Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, 40(5), 255–271.
- Leenders, R. T. A. J. (2002). Modelling social influence through network autocorrelation: Constructing the weight matrix. *Social Networks*, 24(1), 21–47.
- Lu, X., White, H. (2014). Robustness checks and robustness tests in applied economics. *Journal of Econometrics*, 178(P1), 194–206.
- Leduc, S., Wilson, D. (2015). Are state governments roadblocks to federal stimulus? Evidence on the flypaper effect of highway grants in the 2009 Recovery Act. Federal Reserve Bank of San Francisco Working Paper Series, 2013-16.
- Lundberg, J. (2006). Spatial interaction model of spillovers from locally provided public services. *Regional Studies*, 40(6), 631–644.
- Lundberg, J. (2014). On the definition of W in empirical models of yardstick competition. *Annals of Regional Science*, 52(2), 597–610.
- MacKinnon, J. G. (1983). Model specification tests against non-nested alternatives. *Econometric Reviews*, 2(1), 85–110.
- MacKinnon, J. G. (1992). Model specification tests and artificial regressions. *Journal of Economic Literature*, 30(1), 102–146.
- Matsuki, S. (2010). *The twelve months of municipal budgeting [Jichitai zaimu no 12-kagetsu]*, Tokyo: Gakuyo Shobo. [In Japanese]
- Ministry of Internal Affairs and Communications (2014a). *Annual Report on Municipal Finances [Shi-Cho-Son betsu Kessan Joukyo Shirabe]*. http://www.soumu.go.jp/iken/kessan_jokyo_2.html (Accessed 16 September 2014). [In Japanese]
- Ministry of Internal Affairs and Communications (2014b). *The Fiscal Indices Tables for Similar Municipalities [Ruiji Dantai Betsu Shichoson Zaisei Shisu Hyo]*. <http://www.soumu.go.jp/iken/ruiji/> Accessed 16 September 2014. [In Japanese]
- Murdoch, J. C., Rahmatian, M., Thayer, M. A. (1993). A spatially autoregressive median voter model of recreational expenditures. *Public Finance Quarterly*, 21(3), 334–350.
- Negishi, M. (2007). The role of modern fiscal analysis: On the historical processes leading to the establishment of the Fiscal Index Tables of Similar Municipalities [*Gendai no chihozaisei bunseki no yakuwari: Ruijidantaibestu Shichoson Zaisei Shisuhyo no seiritsu katei no bunseki wo chushin to shite*], *Local Public Finance [Chiho Zaisei]*, 46(9), 273–293. [In Japanese]
- Nikkei Digital Media (2014). Regional information files. Nikkei Economic Electronic Databank System. <http://www.nikkei.co.jp/needs/>. Accessed (Accessed 16 September 2014).
- Nishihama, S. (2007). Fiscal analysis of expenditure structures among similar municipalities [*Saishutsu Kozo kara Mita Ruiji Dantai no Zaisei Bunseki*], *Jichi Osaka*, 57(12), 31–38. [In Japanese]
- Nogare, C. D., Galizzi, M. M. (2011). The political economy of cultural spending: Evidence from Italian cities. *Journal of Cultural Economics*, 35(3), 203–231.
- Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), 967–1012.
- Pesaran, M. H., Weeks, M. (2003). Non-nested hypothesis testing: An overview. In B. H. Baltagi (Ed.), *A Companion to Theoretical Econometrics* (Chapter 13). Texas: Blackwell.

- Pollak, R. A., Wales, T. J. (1991). The likelihood dominance criterion: A new approach to model selection. *Journal of Econometrics*, 47(2–3), 227–242.
- Qu, X., Lee, L. F. (2015). Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics*, 184(2), 209–232.
- Revelli, F. (2001). Spatial patterns in local taxation: tax mimicking or error mimicking? *Applied Economics*, 33(9), 1101–1107.
- Revelli, F. (2003). Reaction or interaction? Spatial process identification in multi-tiered government structures. *Journal of Urban Economics*, 53(1), 29–53.
- Revelli, F. (2006). Performance rating and yardstick competition in social service provision. *Journal of Public Economics*, 90(3), 459–475.
- Revelli, F., Tovmo, P. (2007). Revealed yardstick competition: Local government efficiency patterns in Norway. *Journal of Urban Economics*, 90(3), 459–475.
- Rincke, J. (2010). A commuting-based refinement of the contiguity matrix for spatial models, and an application to local police expenditures. *Regional Science and Urban Economics*, 40(5), 324–330.
- Saha, A., Shumway, R., Talpaz, H. (1994). Performance of likelihood dominance and other nonnested model selection criteria: Some Monte Carlo results. Department of Agricultural Economics Working Paper, College Station: Texas A&M University.
- Salmon P. (1987). Decentralization as an incentive scheme. *Oxford Review of Economic Policy*, 3(2), 97–117.
- Sandler, T. (1992). *Collective action: Theory and applications*. Ann Arbor: University of Michigan Press.
- Shleifer, A. (1985). A theory of yardstick competition. *RAND Journal of Economics*, 16(3), 319–327.
- Smith, T. (2009). Estimation bias in spatial models with strongly connected weight matrices. *Geographical Analysis*, 41(3), 307–332.
- Solé-Ollé, A. (2006). Expenditure spillovers and fiscal interactions: Empirical evidence from local governments in Spain. *Journal of Urban Economics*, 59(1), 32–53.
- Stakhovych, S., Bijmolt, T.H.A. (2008). Specification of spatial models: A simulation study on weights matrices. *Papers in Regional Science*, 88(2), 389–408.
- Stetzer, F. (1982). Specifying weights in spatial forecasting models: The results of some experiments. *Environment and Planning A*, 14(5), 571–584.
- Werck, K., Heydels, B., Geys, B. (2008). The impact of “central places” on spatial spending patterns: Evidence from Flemish local government. *Journal of Cultural Economics*, 32(1), 35–58.
- Wildasin, D. E. (1988). Nash equilibria in models of fiscal competition. *Journal of Public Economics*, 35(2), 229–240.
- Williams, A. (1966). The optimal provision of public goods in a system of local government. *Journal of Political Economy*, 74(1), 18–33.
- Wrede, M. (2001). Yardstick competition to tame the Leviathan. *European Journal of Political Economy*, 17(4), 705–721.
- Yu, J., de Jong, R., Lee, L. F. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*, 146(1), 118–134.

Table 1. Descriptive statistics

Variable	Year(s)	Mean	S.D.	Min.	Max.	Obs.	Unit
Per capita expenditure ^a	All	55.0	34.9	19.6	359.6	4,911	0,000 yen
	2008	50.6	30.1	19.6	338.0	1,637	
	2009	56.3	36.0	20.9	358.8	1,637	
	2010	58.2	37.8	22.4	359.6	1,637	
Per capita income ^b	All	114.8	38.0	45.4	670.0	4,911	0,000 yen
	2008	118.6	40.1	45.4	670.0	1,637	
	2009	116.5	38.5	45.7	632.1	1,637	
	2010	109.2	34.6	45.6	548.0	1,637	
Per capita grants ^a	All	19.5	20.7	0.0	167.1	4,911	0,000 yen
	2008	18.4	19.7	0.0	155.6	1,637	
	2009	19.2	20.4	0.0	160.3	1,637	
	2010	20.9	21.8	0.0	167.1	1,637	
Population ^a	All	71.3	181.1	0.5	3,620.6	4,911	000 persons
	2008	71.2	180.4	0.5	3585.8	1,637	
	2009	71.3	181.2	0.5	3606.0	1,637	
	2010	71.3	181.8	0.5	3620.6	1,637	
Surface area ^a	All	210.8	242.8	3.5	2,177.7	4,911	km ²
	2008	210.7	242.8	3.5	2177.7	1,637	
	2009	210.8	242.8	3.5	2177.7	1,637	
	2010	210.9	242.9	3.5	2177.7	1,637	
Proportion of young ^a	All	12.8	2.2	3.7	21.8	4,911	Percent
	2008	12.9	2.2	4.6	21.8	1,637	
	2009	12.8	2.2	4.0	21.4	1,637	
	2010	12.6	2.3	3.7	21.4	1,637	
Proportion of elderly ^a	All	26.8	6.8	10.7	56.9	4,911	Percent
	2008	26.3	6.8	10.7	55.2	1,637	
	2009	26.9	6.8	11.3	56.5	1,637	
	2010	27.4	6.8	11.8	56.9	1,637	

Sources. *a* – *Annual Report on Municipal Finances [Shi-Cho-Son betsu Kessan Joukyo Shirabe]* for relevant years (Ministry of Internal Affairs and Communication 2014a); *b* – Regional Information Files, Nikkei Economic Electronic Databank System (Nikkei Digital Media 2014).

Table 2. Estimation results

	Model FI	Model FC	Model FF	Model II	Model IC	Model IF	Model CI	Model CC	Model CF
W: Weights for interaction	FM	FM	FM	ID	ID	ID	CG	CG	CG
M: Weights for error term	ID	CG	FM	ID	CG	FM	ID	CG	FM
Fiscal interaction (ρ)	0.439*** (0.038)	0.437*** (0.038)	0.434*** (0.051)	0.596*** (0.143)	0.641*** (0.126)	0.582*** (0.125)	0.081*** (0.027)	0.456*** (0.035)	0.093*** (0.022)
Error interaction (λ)	0.708*** (0.137)	0.072*** (0.025)	0.035 (0.122)	0.511*** (0.187)	0.045* (0.026)	0.510*** (0.048)	0.567*** (0.189)	-0.457*** (0.051)	0.515*** (0.047)
Per capita income	0.062* (0.033)	0.071** (0.032)	0.070** (0.032)	0.070** (0.034)	0.061* (0.033)	0.067** (0.033)	0.092*** (0.033)	0.034 (0.025)	0.095*** (0.033)
Per capita grants	1.768*** (0.126)	1.774*** (0.125)	1.809*** (0.127)	2.356*** (0.115)	2.335*** (0.113)	1.845*** (0.138)	2.398*** (0.114)	1.925*** (0.112)	1.892*** (0.136)
Population	0.111 (0.135)	0.096 (0.133)	0.093 (0.135)	0.207 (0.138)	0.218 (0.136)	0.078 (0.158)	0.166 (0.137)	0.236** (0.110)	0.013 (0.157)
Surface area	-0.054 (0.088)	-0.060 (0.090)	-0.066 (0.088)	-0.070 (0.090)	-0.071 (0.090)	-0.049 (0.088)	-0.075 (0.090)	-0.072 (0.070)	-0.061 (0.088)
Proportion young	-1.138 (0.805)	-0.990 (0.798)	-0.829 (0.791)	-1.248 (0.815)	-1.211 (0.808)	-1.254 (0.801)	-1.178 (0.818)	-0.657 (0.700)	-1.052 (0.800)
Proportion elderly	-0.570 (0.542)	-0.581 (0.518)	-0.622 (0.511)	-1.132** (0.540)	-1.044** (0.519)	-0.936* (0.525)	-1.174** (0.543)	-0.819** (0.399)	-0.925* (0.525)
var(ϵ)	55.260*** (1.369)	55.500*** (1.374)	55.707*** (1.379)	57.331*** (1.418)	57.412*** (1.420)	55.709*** (1.382)	57.385*** (1.419)	51.932*** (1.564)	55.716*** (1.383)
Log Likelihood	-11,227	-11,232	-11,236	-11,278	-11,280	-11,243	-11,280	-11,261	-11,244
AIC	22,476	22,486	22,494	22,578	22,581	22,508	22,581	22,544	22,511
BIC	22,543	22,553	22,561	22,645	22,648	22,575	22,648	22,611	22,578
Sample size	3,274	3,274	3,274	3,274	3,274	3,274	3,274	3,274	3,274

Notes: (i) ***: $p \leq .01$; **: $.01 < p \leq .05$; *: $.05 < p \leq .10$. (ii) FM, ID, and CG refer respectively to spatial weights based on FITS-M, inverse distance, and contiguity. (iii) Standard errors are in parentheses. (iv) All the models allow for fixed municipal effects as well as fixed time effects.

Table 3. Estimation results with the hybrid weights matrix

	Model HI	Model HC	Model HF	Model HH	Model FH	Model IH	Model CH
W: Weights for interaction	HB	HB	HB	HB	FM	ID	CG
M: Weights for error term	ID	CG	FM	HB	HB	HB	HB
Fiscal interaction (ρ)	0.389*** (0.032)	0.402*** (0.032)	0.377*** (0.039)	0.434*** (0.045)	0.368*** (0.048)	0.448*** (0.146)	0.038 (0.025)
Error interaction (λ)	0.300 (0.204)	0.001 (0.026)	0.119 (0.105)	-0.067 (0.076)	0.263*** (0.044)	0.392*** (0.039)	0.404*** (0.040)
Per capita income	0.069** (0.032)	0.068** (0.031)	0.074** (0.032)	0.062** (0.031)	0.084** (0.033)	0.086** (0.034)	0.105*** (0.034)
Per capita grants	1.833*** (0.119)	1.822*** (0.118)	1.826*** (0.120)	1.780*** (0.126)	1.822*** (0.128)	2.003*** (0.132)	2.051*** (0.132)
Population	0.115 (0.132)	0.111 (0.130)	0.088 (0.138)	0.116 (0.125)	0.065 (0.152)	0.105 (0.165)	0.049 (0.165)
Surface area	-0.063 (0.087)	-0.067 (0.087)	-0.065 (0.087)	-0.066 (0.086)	-0.066 (0.089)	-0.066 (0.090)	-0.073 (0.090)
Proportion young	-1.128 (0.792)	-1.037 (0.784)	-1.055 (0.786)	-0.969 (0.780)	-1.141 (0.801)	-1.386* (0.808)	-1.308 (0.808)
Proportion elderly	-0.549 (0.516)	-0.536 (0.499)	-0.597 (0.508)	-0.508 (0.490)	-0.591 (0.533)	-0.759 (0.549)	-0.783 (0.550)
var(ϵ)	54.654*** (1.357)	54.654*** (1.357)	54.716*** (1.359)	54.481*** (1.369)	54.939*** (1.362)	55.256*** (1.374)	55.338*** (1.378)
Log Likelihood	-11,217	-11,218	-11,218	-11,218	-11,219	-11,236	-11,240
AIC	22,457	22,459	22,458	22,458	22,461	22,495	22,501
BIC	22,524	22,526	22,525	22,525	22,528	22,562	22,568
Sample size	3,274	3,274	3,274	3,274	3,274	3,274	3,274

Notes: (i) ***: $p \leq .01$; **: $.01 < p \leq .05$; *: $.05 < p \leq .10$. (ii) FM, ID, and CG refer respectively to spatial weights based on FITS-M, inverse distance, and contiguity. (iii) Standard errors are in parentheses. (iv) All the models allow for fixed municipal effects as well as fixed time effects.

Table 4. Estimation results with the hybrid weights matrix year-regional dummies

	Model HI'	Model HC'	Model HF'	Model HH'	Model FI'	Model FC'	Model FF'	Model FH'
W: Weights for interaction	HB	HB	HB	HB	FM	FM	FM	FM
M: Weights for error term	ID	CG	FM	HB	ID	CG	FM	HB
Fiscal interaction (ρ)	0.389*** (0.032)	0.386*** (0.032)	0.366*** (0.041)	0.439*** (0.044)	0.438*** (0.038)	0.436*** (0.038)	0.460*** (0.048)	0.380*** (0.047)
Error interaction (λ)	-0.158 (0.265)	-0.010 (0.026)	0.087 (0.123)	-0.120 (0.076)	0.423** (0.206)	0.053** (0.025)	-0.086 (0.149)	0.223*** (0.046)
Per capita income	0.042 (0.033)	0.043 (0.033)	0.047 (0.034)	0.034 (0.033)	0.033 (0.034)	0.032 (0.034)	0.024 (0.034)	0.048 (0.035)
Per capita grants	1.817*** (0.121)	1.816*** (0.121)	1.811*** (0.123)	1.751*** (0.126)	1.724*** (0.128)	1.727*** (0.128)	1.733*** (0.127)	1.761*** (0.130)
Population	0.127 (0.132)	0.128 (0.133)	0.113 (0.139)	0.130 (0.124)	0.121 (0.135)	0.120 (0.135)	0.133 (0.131)	0.101 (0.151)
Surface area	-0.051 (0.087)	-0.051 (0.087)	-0.049 (0.087)	-0.049 (0.086)	-0.042 (0.088)	-0.041 (0.089)	-0.045 (0.087)	-0.047 (0.089)
Proportion young	-1.323* (0.796)	-1.331* (0.797)	-1.368* (0.800)	-1.246 (0.792)	-1.366* (0.808)	-1.378* (0.808)	-1.257 (0.801)	-1.442* (0.808)
Proportion elderly	-0.641 (0.539)	-0.648 (0.542)	-0.674 (0.549)	-0.612 (0.531)	-0.665 (0.559)	-0.645 (0.556)	-0.624 (0.547)	-0.613 (0.564)
var(ϵ)	54.301*** (1.349)	54.316*** (1.349)	54.382*** (1.351)	53.995*** (1.360)	54.954*** (1.361)	54.952*** (1.360)	55.001*** (1.364)	54.564*** (1.352)
Log Likelihood	-11,206	-11,206	-11,206	-11,205	-11,215	-11,215	-11,217	-11,206
AIC	22,463	22,463	22,462	22,461	22,480	22,479	22,484	22,462
BIC	22,615	22,615	22,615	22,613	22,633	22,632	22,636	22,614
Sample size	3,274	3,274	3,274	3,274	3,274	3,274	3,274	3,274

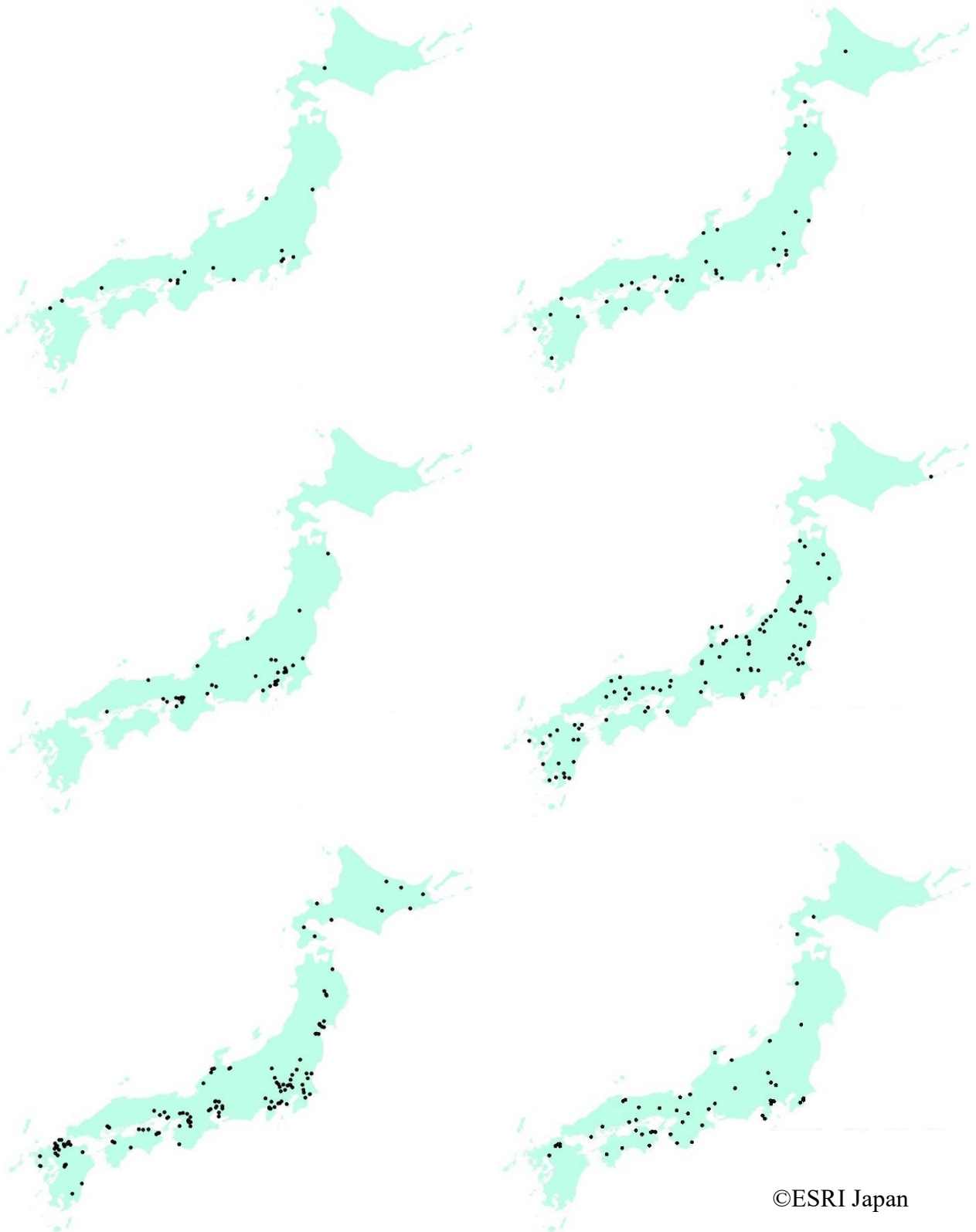
Notes: (i) ***: $p \leq .01$; **: $.01 < p \leq .05$; *: $.05 < p \leq .10$. (ii) FM, ID, and CG refer respectively to spatial weights based on FITS-M, inverse distance, and contiguity. (iii) Standard errors are in parentheses. (iv) All the models allow for fixed municipal effects as well as regionally differentiated temporal effects (the interactions between time dummies and regional dummies).

Table 4 (Continued). Estimation results with the hybrid weights matrix year-regional dummies

	Model II'	Model IC'	Model IF'	Model IH'	Model CI'	Model CC'	Model CF'	Model CH'
W: Weights for interaction	ID	ID	ID	ID	CG	CG	CG	CG
M: Weights for error term	ID	CG	FM	HB	ID	CG	FM	HB
Fiscal interaction (ρ)	0.479** (0.199)	0.444** (0.197)	0.466*** (0.172)	0.186 (0.203)	0.096*** (0.029)	0.439*** (0.037)	0.077*** (0.023)	0.025 (0.025)
Error interaction (λ)	0.269 (0.261)	0.044 (0.028)	0.525*** (0.047)	0.389*** (0.040)	0.026 (0.312)	-0.449*** (0.053)	0.523*** (0.048)	0.385*** (0.041)
Per capita income	0.059* (0.035)	0.060* (0.035)	0.058* (0.035)	0.074** (0.035)	0.061* (0.034)	0.034 (0.028)	0.061* (0.035)	0.076** (0.035)
Per capita grants	2.367*** (0.116)	2.365*** (0.116)	1.832*** (0.140)	2.005*** (0.135)	2.352*** (0.115)	1.987*** (0.113)	1.823*** (0.140)	2.010*** (0.135)
Population	0.192 (0.138)	0.188 (0.138)	0.054 (0.161)	0.096 (0.167)	0.176 (0.136)	0.225* (0.115)	0.037 (0.161)	0.088 (0.166)
Surface area	-0.063 (0.089)	-0.062 (0.090)	-0.045 (0.088)	-0.056 (0.090)	-0.062 (0.089)	-0.062 (0.071)	-0.043 (0.088)	-0.055 (0.090)
Proportion young	-1.383* (0.823)	-1.409* (0.823)	-1.394* (0.812)	-1.562* (0.815)	-1.383* (0.818)	-0.758 (0.729)	-1.411* (0.811)	-1.568* (0.815)
Proportion elderly	-1.333** (0.562)	-1.317** (0.562)	-0.904 (0.572)	-0.833 (0.580)	-1.298** (0.555)	-0.986** (0.454)	-0.882 (0.572)	-0.840 (0.579)
var(ϵ)	57.117*** (1.412)	57.092*** (1.412)	55.352*** (1.374)	55.062*** (1.370)	56.971*** (1.410)	51.781*** (1.563)	55.263*** (1.372)	55.073*** (1.371)
Log Likelihood	-11,269	-11,269	-11,233	-11,229	-11,266	-11,247	-11,231	-11,229
AIC	22,589	22,587	22,516	22,508	22,583	22,545	22,511	22,508
BIC	22,741	22,740	22,668	22,661	22,735	22,697	22,663	22,661
Sample size	3,274	3,274	3,274	3,274	3,274	3,274	3,274	3,274

Notes: (i) ***: $p \leq .01$; **: $.01 < p \leq .05$; *: $.05 < p \leq .10$. (ii) FM, ID, and CG refer respectively to spatial weights based on FITS-M, inverse distance, and contiguity. (iii) Standard errors are in parentheses. (iv) All the models allow for fixed municipal effects as well as regionally differentiated temporal effects (the interactions between time dummies and regional dummies).

Figure 1. Selected groups of “similar” municipalities.



Notes: The panels pick up six categories of “similar municipalities” and plot them on a map of Japan with ESRI Japan (2011).