Working Paper Series (E)

No.27

Identifying Tax Mimicking in Municipal Health Insurance: Evidence from A Boundary Reform

Michihito Ando
Reo Takaku

July 2016
http://www.ipss.go.jp/publication/e/WP/IPSS_WPE27.pdf

National Institute of Population and Social Security Research
Hibiya Kokusai Building 6F, 2-2-3 Uchisaiwaicyo, Chiyoda-ku, Tokyo 100-0011
http://www.ipss.go.jp

Identifying Tax Mimicking in Municipal Health Insurance: Evidence from A Boundary Reform*

Michihito Ando† Reo Takaku‡

July 1, 2016

Abstract

This paper examines whether tax mimicking occurs among municipalities in the setting of health insurance tax levels in Japan. To uncover the effects of strategic mimicking behavior among neighboring municipalities, we exploit the fact that insurance tax levels sharply dropped when municipalities experienced municipal amalgamation during the Great Heisei Amalgamation, which took place during the mid-2000s. Utilizing the incidence of neighbor amalgamation as an instrumental variable, we investigate how insurance tax levels in neighbor municipalities affect tax levels in non-amalgamated municipalities. Results suggest that there has been significant mimicking behavior in non-amalgamated municipalities whose insurance tax levels were higher than those of their neighbors before the Great Heisei Amalgamation. The other non-amalgamated municipalities have not responded to changes in neighbor insurance tax levels. We also discuss the issues of internal and external validity in the identification of tax mimicking.

JEL classification: H20, H71, H77
Keywords: strategic interaction, tax mimicking, quasi-experiment, instrumental variable, municipal amalgamation

*We are grateful to Shun-ichiro Bessho, Masayoshi Hayashi, Che-Yuan Liang, Katsuyoshi Nakazawa and seminar participants at Aoyama Gakuin University, Hitotsubashi University, University of Tokyo, Gakushuin University, Tohoku University, and Nagoya University for their comments and suggestions. The views expressed in this paper do not necessarily reflect those of the National Institute of Population and Social Security Research or the Institute of Health Economics and Policy. This work was supported by JSPS KAKENHI Grant Number JP15K17085 and Health Labour Sciences Research Grant (MHLW Grant) [Grant Number H25-Seisaku-Ippan-004]. All errors are our own.

†National Institute of Population and Social Security Research, Hibiya Kokusai Building 6th Floor, 2-2-3 Uchisaiwaicho, Chiyoda-ku, Tokyo, 100-0011, Japan, +81-3-3595-2984, andou-michihito@ipss.go.jp

‡Institute for Health Economics and Policy, No.11 Toyo-kaiji Bldg, 1-5-11, Nishi-Shinbashi, Minato-ku, Tokyo, 105-0003, Japan, reo.takaku@ihep.jp

1 Introduction

Since the seminal work of Besley and Case (1995) numerous empirical studies have been published on the fiscal interaction between local governments within a variety of policy contexts. On the literature of tax competition, Allers and Elhorst (2005) summarize previous findings and conclude that tax competition exists with a neighborhood effect of between 0.2 and 0.6. More recently Costa-Font et al. (2014) provide a meta-regression analysis on tax competition and also find robust evidence of tax interaction. Most existing studies that are referred to by Allers and Elhorst (2005) and Costa-Font et al. (2014) rely on spatial econometric approaches such as Spatial Lag (SL) models and Spatial Instrumental Variables (SIV) models.

On the other hand, some recent quasi-experimental studies have made the criticism that causal interpretation of spatial econometric models is often implausible because estimation with these models is not based on clear identification strategies (Gibbons and Overman 2012). Inspired by this sort of argument, there is an increasing number of studies that examine the strategic interaction of local governments with careful identification strategies that do not rely on spatial econometric identification methods.

For example, Lyytikäinen (2012) employs an instrumental variables approach that is based on a policy-based quasi-experiment with a first-differenced model. He shows there is no strategic interaction in the setting of local property taxes in Finland. Baskaran (2014) examines the existence of tax competition in German municipalities with a similar first-differenced model to that of Lyytikäinen (2012), exploiting the exogenous state-level policy changes as an instrument. He also finds no evidence of tax competition. Isen (2014) utilizes a regression discontinuity design that is based on close elections in Ohio. He also finds no evidence of spatial competition between counties, municipalities and school districts. Parchet (2014), like Baskaran (2014), uses state-level fiscal reforms as an exogenous source of variation in tax rates of local governments in Switzerland. He finds negative strategic interaction in personal income tax rates, which implies that the tax rates are strategic substitutes rather than strategic complements.[1]

Following this quasi-experimental literature, this paper examines the strategic

[1] Other recent related quasi-experimental studies on tax competition include Eugster and Parchet (2014) and Agrawal (2015), both of which investigate tax competition by investigating spatial gradients in local taxes across state borders without estimating tax reaction functions.

interaction of tax levels in a municipal public health insurance scheme in Japan called "Citizens' Health Insurance" (CHI, *Kokumin Kenkoh Hoken*), exploiting an explicit source of exogenous variation. The identifying variation of this paper comes from large-scale municipal amalgamations implemented in the 2000s, the "Great Heisei Amalgamation." Because the CHI's insurance tax levels tend to decrease sharply in amalgamated municipalities in comparison to their non-amalgamated counterparts for several reasons, this sharp reduction in neighbor insurance tax levels can be exploited as an identifying variation for the effect of the changes in neighbor insurance tax levels on own insurance tax levels.

Our main contribution to the literature is summarized as follows. Contrary to the findings of a few recent quasi-experimental studies, we provide plausible evidence of tax mimicking in a subgroup of municipalities. That is, combining a difference-in-differences (DID) method and an instrumental variables (IV) approach, we find that a sharp reduction in CHI insurance taxes in neighboring amalgamated municipalities leads to significant reduction in insurance taxes in *some* non-amalgamated municipalities.

More specifically, we find that only municipalities that had higher insurance tax levels than their neighbors in pre-treatment periods decreased their tax levels when they were faced with a reduction in tax levels in neighbor municipalities. On the other hand, we do not find any evidence that the other municipalities in our sample responded to reduction in neighbor insurance tax levels. This result seems plausible because municipalities in which tax levels are already lower than their neighbors should have a weaker incentive to decrease their tax levels when their neighbors do so.

We also contribute to the literature by critically addressing the issues of the internal and external validity of using a quasi-experimental variation for the identification of tax mimicking. First, when it comes to internal validity, the previous studies mentioned above emphasize that they use seemingly exogenous variations to overcome identification problems in the conventional spatial econometrics method. We argue, however, that the advantage of quasi-experimental methods over conventional spatial econometric methods is somewhat unclear, in particular when the intensity of exploited "exogenous" variation is spatially correlated. In this case, the utilized variation may be "external" in the sense that the variation is caused by the central government such as Lyytikäinen (2012) or an upper-level government (i.e. state) such as Baskaran (2014), but may not be "exogenous" (i.e. violating exclusion restriction in the context of IV strategy).[2]

In our case, we discuss the fact that our exogeneity assumption may also be violated by both "spatial selection into treatment" and "censoring by death". [3] We then provide placebo tests utilizing pre-treatment outcomes to check whether such violations undermine our identification and estimation. Lyytikäinen (2012) also adopts a similar placebo test, but we use data with a much longer pre-treatment period to verify our identifying assumptions.

Second, external validity is also an important issue, particularly because some previous quasi-experimental studies use very specific variations for identification and also use subgroups of municipalities for estimation. For example, Lyytikäinen (2012) exploits neighboring municipalities' reform-induced increases in property tax rates to construct an instrument. Because the national-level tax reform that he exploits can affect the tax rates of the municipalities in question as well, Lyytikainen also conducts analysis controlling for reform-induced "own imposed increases" or using a subsample of municipalities in which pre-reform tax rates are relatively high and not directly affected by the reform. Baskaran (2014) and Parchet (2014) examine strategic interaction between neighboring municipalities that belong to different states. Isen (2014)'s identification strategy is a regression discontinuity design exploiting close referenda on tax increases in neighboring municipalities.

Contrary to most studies with conventional spatial econometric approaches, these four studies do not find tax mimicking behavior: Lyytikäinen (2012), Baskaran (2014) and Isen (2014) find no strategic interaction and Parchet (2014) observes negative interaction. The internal validity of their findings seem to be more plausible than previous studies in the sense that their identifying variations are more exogenous than those of their predecessors. There is, however, a possibility that they find no mimicking behavior because they investigate specific cases where mimicking behavior of local governments is unlikely to happen.[4] In fact, these studies stress

[2]Heckman (2000) and Deaton (2010) explain the distinction between the terms "external" and "exogenous", although Deaton (2010) notes that these terms are "hardly standard".

[3]The problem of "censoring by death" is addressed by, among others, Frangakis and Rubin (2002) and Rubin (2006). See Section 6.1 for possible bias caused by "censoring by death" in the context of our research.

[4]Lyytikäinen (2012) essentially captures the specific response of municipalities with *higher* pre-reform tax rates to tax increases in municipalities with *lower* pre-reform tax rates. Baskaran (2014) and Parchet (2014) investigate tax mimicking behavior of municipalities *across* the border of different states, not *within* the same state where mimicking

that their results may be derived from the specific sources of exogenous variation or institutional settings in their research designs.

Our study is also not free from threats to its external validity. Our identifying variation is local and specific and we use a subsample of non-amalgamated municipalities for analysis. Contrary to the previous quasi-experimental studies, we nonetheless find quite strong mimicking behavior for the subgroup of municipalities with higher pre-reform tax levels. In addition, the estimated effect is larger than a counterpart estimate obtained by utilizing a conventional spatial instrument. Overall, our estimation results suggest that local governments respond sensitively to their neighbors when they have incentives to do so.

The rest of this paper proceeds as follows. Section 2 provides a brief explanation of the relevant institutional background. Section 3 explains our research design. In section 4 we explain our dataset and data construction and then provide some graphical analysis. Section 5 shows the baseline results of this study and also presents some robustness checks. Section 6 discusses potential threats to the validity of our identification and provides placebo tests to defend our primary finding. In Section 7 we investigate how tax mimicking occurs and compare our main results with estimation results obtained using a spatial-IV approach. Section 8 concludes.

2 Background

2.1 Citizen's Health Insurance (CHI)

Almost all residents of Japan are entitled to join some form of public insurance scheme, and the payment scheme is uniformly set by the central government. Under this universal health insurance system, people can in principle visit any hospital or clinic to receive medical treatment; this is often called a "free access" scheme. There is no systematic gatekeeper scheme, and insurers, governments and hospital/clinics do not have explicit measures to control patient visits except for co-payment schemes, which are in principle determined by the central government.

behavior is more likely to happen. Isen (2014) also estimates somewhat specific mimicking behavior because what is identified is the effect of neighbor tax increases realized by close elections observed in the dataset of local referenda, which seem to be rather restrictive circumstances. In addition, as the author points out, the conventional story of yardstick competition may be less relevant in Ohio because tax referenda prevent politicians from directly determining tax levels.

Whereas the entire universal health insurance system is strongly regulated by the central government, public health insurers are deeply fragmented with over 3000 insurers. Largely, they are divided into three categories: employment-based health insurance, municipal health insurance, and the Health Program for the Elderly. Under the current system, people under 75 years old are enrolled in employment-based health insurance if they or the head of their household work for medium or large-sized firms. Employment-based health insurance covers 65 percent of the population under 75 years old. Otherwise, people under age 75 are enrolled in a municipal health insurance program called Citizen's Health Insurance (CHI). CHI mainly serves people who are self-employed, retired but under 75 years old, or employed by a small business and their dependents. The Health Program for the Elderly, which has been reformed over the years, serves people aged 75 years and older.

The health expenditures in CHI are financed as follows.

$$\begin{aligned} CHI\ health\ expenditure &+ Other\ expenditures \\ &\simeq Insurance\ tax\ revenues + Statutory\ transfers \\ &+ Discretionary\ transfers + Other\ transfers \end{aligned}$$

In the left-hand side of the above equation, CHI health expenditure is determined by the healthcare utilization of the CHI insured. Other expenditures consist mainly of fiscal transfers to the Health Program for the Elderly. On the right-hand side, insurance tax revenues are tax (or premium) revenues from those insured by the CHI.[5] Statutory transfers consist of matching transfers from municipalities, prefectures and the central government and equalization transfers from the central government. Discretionary transfers are mostly non-statutory transfers that CHI insurers finance from their municipal general budgets and CHI funds, which are meant to suppress the insurance tax burdens of those insured by CHI or stabilize CHI accounts. Because CHI insurers, mostly municipalities, cannot directly control most health expenditures and transfers except for discretionary transfers, discretionary transfers are in most cases the only effective means to suppress insurance tax levels.

[5]Some municipalities levy insurance *taxes* for CHI while the others charge insurance *premiums*, but we simply refer to both of them as insurance *taxes* because there is no substantive difference between them as far as this paper is concerned.

When municipalities impose the insurance tax on those enrolled in CHI, they can choose the tax bases from which they raise the insurance tax. In principle, municipalities can use four bases: per enrollee (*kintou wari*), per household (*setai wari*), household income (*shotoku wari*) and assets (*shisan wari*). The per-enrollee base is a head count amount that is levied on all enrollees in CHI, while the per-household base is a fixed amount for all households. These two parts of the insurance tax are in fact lump-sum taxes in the sense that the insurance tax levels are the same regardless of individual or household income levels. The other two bases follow the "ability to pay" principle. That is, income-based and asset-based components are levied in proportion to household income and fixed assets respectively. Hence, the total amount of the insurance tax that a household must pay varies depending on its size, income and the value of its assets.

Since municipalities adopt different combinations of the above four tax bases and their levels, the total tax burden of the insurance tax varies considerably across municipalities even for households with similar sizes, incomes, and assets. Under this system, it is impossible to directly compare the whole level of insurance tax. Nonetheless, it is widely recognized that the regional disparity in CHI tax levels is quite large (Ikegami et al. 2011). Based on a simulation with detailed institutional settings from 2010, Takaku et al. (2014) find that for that year the highest CHI tax levels in some municipalities were three or four times higher than the lowest tax levels in some other municipalities, even for a household with the same income and household structure.

2.2 Tax mimicking in health insurance tax in CHI

From a theoretical perspective, it is likely that a strategic interaction occurs in CHI. One plausible explanation is the yardstick competition (Besley and Case 1995), which suggests that local politicians try to buy votes by adjusting insurance tax levels in CHI to a lower level than insurance tax levels in other reference municipalities, because voters evaluate the performance of their local politicians through such interregional comparisons.[6]

[6] Another possible source of the strategic interaction accrues from tax competition based on the mobility of tax bases (Tiebout 1956). This theory does not seem to be relevant to our case, however, because the pressure of tax competition on municipalities is presumably weak since the typical enrollees of CHI are the elderly and the self-employed who are relatively immobile. See Brülhart and Parchet (2014) who show even high-income

Of course, the reduction of the CHI tax level may be financed by decreases in other local public services or increases in other local tax rates and may not necessarily lead to greater political support. But it is possible that many voters, in particular CHI enrollees, are more keenly aware of CHI tax levels than other indices of local public finance. This is a plausible conjecture because the political process of budgetary planning is far from transparent to voters, whereas CHI enrollees know how much they have to pay for their CHI tax. In addition, because most Japanese municipalities adopt a very homogenous local personal income tax rate of 10% and only a few municipalities have different local personal income tax rates, this makes room for local politicians to buy votes through making the CHI tax level lower rather than reforming other major local taxes.

As mentioned earlier, by using non-statutory discretionary transfers from their own general accounts and CHI funds, municipalities can effectively reduce the amount of revenue that has to be raised through the insurance tax. In fact, this is what has happened since at least the 1990s. The fact that insurance tax levels are suppressed by discretionary municipal transfers is important for our study because it implies municipalities at least have an effective institutional device for tax mimicking behavior.

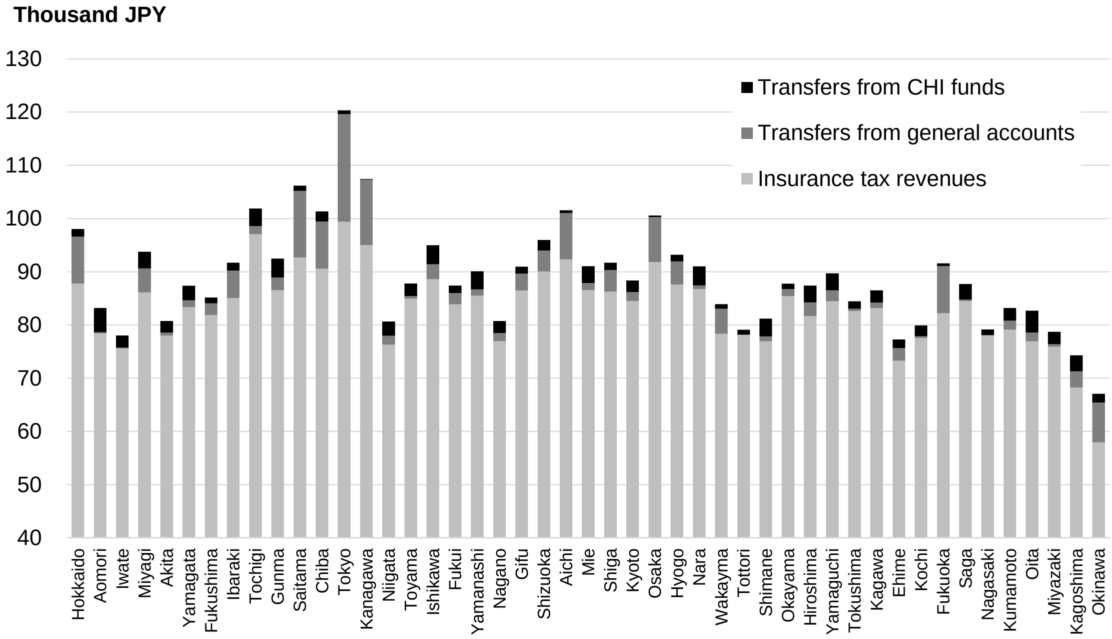
We should also note that the amount of discretionary transfers varies greatly across regions and municipalities. Figure 1 presents the average levels of insurance tax in CHI and discretionary transfers per CHI enrollee at the prefecture level. In this figure, the height of the bar can be loosely interpreted as the amount of insurance tax per enrollee that would be levied if there were no discretionary transfers and if the corresponding revenues were financed by insurance tax. In Tokyo prefecture, for instance, insurance tax per enrollee would be around 120,000 JPY (i.e. around 1,000 USD if 1 USD = 120 JPY) without the discretionary transfers, but the actual tax burden is significantly suppressed by the injection of discretionary transfers; the actual insurance tax level per enrollee is only around 100,000 JPY. This utilization of discretionary transfers suggests that municipalities in Tokyo Prefecture may engage in tax mimicking to make their insurance tax levels lower than neighboring municipalities.

Anecdotal evidence also supports the strategic utilization of discretionary trans-

---

retirees are relatively inelastic with respect to tax rates. In addition, it is unrealistic to assume that municipalities compete to gain more CHI enrollees because typical CHI enrollees do not provide fiscal benefits for municipalities.

Figure 1: Revenues from insurance tax and discretionary transfers per enrollee (2007)



Note: Insurance tax levels are averaged across municipalities (CHI insurers) in each prefecture. Source: Report on Citizens' Health Insurance (Kokumin kenko Hoken No Jittai)

fers. For instance, it is well known that the Japan Communist Party frequently advocates greater transfers from municipal general accounts. In local elections, the candidates of the Communist Party often promise their constituency that they will reduce CHI tax levels through additional injections from general municipal budgets.

3 Empirical strategy

3.1 Neighbor amalgamation as an instrument

A major problem in the identification of tax mimicking behavior is simultaneous and endogenous determination of tax levels across municipalities. To uncover the causal effect of neighboring municipalities' insurance tax levels on a municipality's insurance tax level, we exploit the sharp reduction in insurance tax levels caused by municipal amalgamation during the period of the Great Heisei Amalgamation (2003-2007) as a source of identifying variation. In other words, we use neighbor municipal amalgamation as an instrument that causes a reduction in neighbor insurance tax rates but does not directly affect own insurance tax level.

As is shown in the empirical section, the first-stage relevance of our instrument is testable and we do not fully investigate its mechanism. There are, however, at least three reasons why municipal amalgamation reduces health insurance tax levels. First, the insurance tax level in a newly merged municipality is often set to the lowest found in the old municipalities to avoid an abrupt increase in tax levels for the citizens who lived in the municipality with the lowest pre-amalgamation tax levels. This political incentive seems to be consistent with the theory of political business cycles (Nordhaus 1975; Rogoff 1990), in particular because local politicians are often preparing for a local election in the new municipality. Second, a scale merit can reduce the administrative costs of CHI after amalgamation (Reingewertz 2012). If health care expenditures and operational costs per capita can be reduced as a consequence of amalgamation, it is possible to reduce insurance tax levels. Third, a common pool problem (Weingast et al. 1981) may also have an effect. Before municipal amalgamations, some municipalities may have incentives to increase fiscal transfers from their municipal general account to CHI since the burden of the resulting fiscal imbalance will be shared by the entire population of the new municipal area. (Hinnerich 2009; Jordahl and Liang 2010; Nakazawa 2015)

When it comes to the assumption of exclusion restriction for our instrument, we need to assume that the amalgamation of neighbor municipalities affects own insurance tax level *only through* insurance tax levels in neighbor amalgamated municipalities. We argue that it is hard to come up with other major pathways through which neighbor amalgamation could affect own insurance tax level. This is the case because a municaplity's neighbors' amalgamation is not an incident that is directly related to its own municipal fiscal and healthcare circumstances. Although there could be some minor unobserved pathways that might generate some bias in our IV estimation, the effect of neighbor amalgamation on insurance tax levels, if any, can reasonably be assumed to be caused by a strategic reaction to a merger-induced decrease in neighbor insurance tax levels.

Another concern in our IV strategy is that our exogeneity assumption can be violated because our instrument of neighbor amalgamation is not randomly assigned among municipalities. We therefore adopt a difference-in-differences (DID) approach conditional on observed pre-determined covariates in our IV estimation, exploiting the fact that neighbor amalgamation occurs mostly during 2003-2007 in our sample period. We also implement some placebo tests that reinforce the plausibility of our empirical strategy.

3.2 Great Heisei Amalgamation

To utilize the incident of neighbor amalgamation as a plausible instrument with sufficient variation, we exploit the so called *Great Heisei Amalgamation*, which is a large-scale voluntary municipal amalgamation reform that was carried out in Japan in the 2000's.

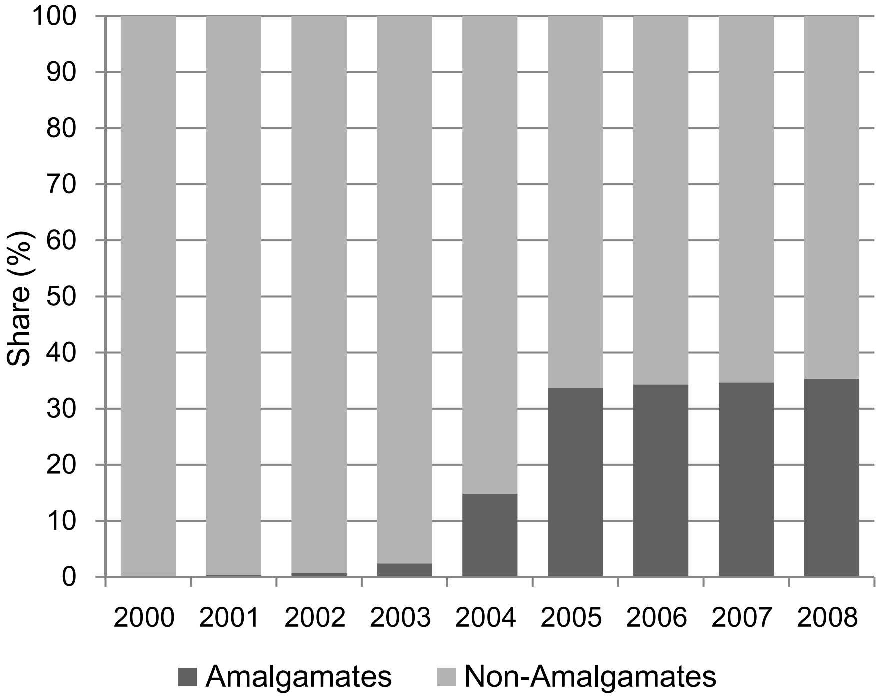
First of all, Japan has three tiers of government: the central government, prefectures, and municipalities. There are 47 prefectures, each of which contains multiple municipalities. Before the Great Heisei Amalgamation there were more than 3,260 municipalities, but this number had decreased to 1750 by 2010. Municipal amalgamation was extensively implemented from 2003 to 2006 because the issuance of special bonds and several other special measures in favor of amalgamation were provided to municipalities that merged during this period.

According to official documents, the intentions behind this drastic amalgamation reform were (1) to strengthen the administrative scale and capability of municipalities for the promotion of decentralization, (2) to enhance the capability of municipalities as the responsible party of social service provision in an era of low birthrates and an aging population, (3) to adapt to people moving greater distances in their daily lives, and (4) to increase administrative efficiency.

To examine the scale of the Great Heisei Amalgamation year by year, Figure 2 presents the share of municipalities that experienced any form of amalgamation during the 2000s. This figure shows that the share of amalgamates increased gradually and reached 35 percent in 2008.

For our empirical strategy, it is noteworthy that municipal amalgamation during the 2000s was voluntary, not compulsory. In order to promote amalgamation, the central government utilized a variety of financial incentives. First, the central government enacted the Devolution of Power Law in 1995. This legislation allowed amalgamated municipalities to issue special subsidized bonds (*Gappei Tokurei Sai*). Although municipal amalgamation required a large amount of special expenditures, municipalities could easily appropriate these costs by issuing these bonds that were almost completely compensated through additional fiscal transfers from the central government. In addition, the central government cut the fiscal equalization grants in the late 1990s and early 2000s and the reduction in fiscal equalization grants was more severe for smaller municipalities. This policy trend may also have encouraged smaller municipalities to amalgamate with neighboring municipalities. We

Figure 2: Share of Municipalities which Experienced Amalgamations



Note: Municipal boundaries are based on fiscal year 2008. “Amalgamates” are the municipalities that experienced any form of amalgamation after 1999 and “Non-amalgamates” are all other municipalities. In this figure, the experience of amalgamation is traced retrospectively based on municipalities in 2008. That is, if town A and town B amalgamated and city C was newly established in 2004, we change the status of city C from non-amalgamated to amalgamated in 2004.

discuss possible threats to our identification strategy arising from these voluntary amalgamations in Section 6.

3.3 IV construction

Because we utilize panel data to control for unobserved fixed effects and amalgamation makes merged municipalities non-identical entities before and after amalgamation, we use the sample of non-amalgamated municipalities for the subsequent analysis.

We then construct our instrument $Z_{-i,t}$ as a population-weighted neighbor amalgamation rate for non-amalgamated municipality i at year t:

$$Z_{-i,t} = \sum_{j \neq i} 1[Merge_{j,t}] w_{i,j}, \tag{1}$$

where

$$w_{i,j} = \frac{1[Neighbor_{i,j,T_0}]Pop_{j,T_0}}{\sum_{j \neq i} 1[Neighbor_{i,j,T_0}]Pop_{j,T_0}}. \tag{2}$$

In equation (1), $1[Merge_{j,t}]$ is an indicator variable that takes one if municipality j experiences merger by year t and otherwise zero. Note that $1[Merge_{j,t}]$ is an indicator of amalgamation that reflects the past experience, not the current event, of amalgamation in order to capture lagged amalgamation effects.[7] In equation (2), $1[Neighbor_{i,j,T_0}]$ is an indicator variable that takes one if municipality i shares a border with municipality j at reference year T_0 and otherwise zero, and Pop_{j,T_0} is the population of municipality j at T_0. T_0 is a pre-amalgamation year used as a reference period. By using this population-weighted indicator, the instrument $Z_{-i,t}$ reflects the intensity of neighbor amalgamations, with an amalgamation of a larger neighbor municipality getting a higher value.

Figure 3 explains our sample choice and variation caused by our instrument using the map of Hokkaido prefecture, which is located in the north of Japan. In this figure, white areas are municipalities that experienced amalgamations during the Great Heisei Amalgamation. These municipalities are excluded from the sample because they do not have the same identities before and after amalgamation. At the same time, their mergers may cause seemingly exogenous suppression in their insurance tax levels. This suppression in neighbor tax levels may then have affected tax levels in non-amalgamated municipalities that share a border with amalgamated municipalities, expressed as darker gray areas. The magnitude of suppression in neighbor tax levels, however, may differ depending on the intensity of neighbor amalgamation, so we use the formula (1) as an instrument that reflects this amalgamation intensity. The lighter gray areas are non-amalgamated municipalities that do not share a border with amalgamated municipalities and therefore have a value of zero for instrument $Z_{-i,t}$.

3.4 Empirical models

Because our instrument of neighbor amalgamation is not randomly assigned, we adopt a difference-in-differences (DID) framework to estimate the magnitude of tax mimicking, controlling for municipality fixed effects.

[7]Amalgamation before 1994, the first year of our sample, is not taken into account in $1[Merge_{j,t}]$.

Figure 3: Amalgamations in Hokkaido during the Great Heisei Amalgamation

Note: White areas are municipalities that experienced amalgamations during the Great Heisei Amalgamation. Darker gray areas are non-amalgamated municipalities that share a border with amalgamated municipalities. Lighter gray areas are non-amalgamated municipalities that do not share a border with amalgamated municipalities. Two mergers indicated with blue lines generated two new municipalities that have enclaves.

First, our baseline model to be estimated is expressed as follows:

$$Y_{i,t} = \alpha_i + \beta_t + \tau Y_{-i,t} + \varepsilon_{i,t}, \tag{3}$$

where $Y_{i,t}$ is the insurance tax level of municipality i at year t, α_i is an individual fixed effect for i, and β_t is a time fixed effect at year t. $Y_{-i,t}$ is the population-weighted average of the insurance tax levels of neighbor municipalities in which the population weight $w_{i,j}$ is used for averaging. τ_t is the parameter of interest and represents the causal effect of $Y_{-i,t}$ on $Y_{i,t}$ at year t.

Second, we also estimate the following model with additional covariates:

$$Y_{i,t} = \alpha_i + \beta_t + \tau Y_{-i,t} + \mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t + \mathbf{X}'_{-i,T_0} \cdot \boldsymbol{\pi}_t + \varepsilon_{i,t}, \tag{4}$$

where $\mathbf{X}_{i,T_0}$ and $\mathbf{X}_{-i,T_0}$ are the vectors of pre-determined covariates (or "initial conditions") of own and neighbor municipalities for the (last) pre-treatment year T_0 and $\boldsymbol{\mu_t}$ and $\boldsymbol{\mu_t}$ are the vectors of time-varying coefficients of these covariates.[8]

[8]We do not use the standard way of introducing time-varying covariates with constant

For $\mathbf{X}_{-i,T_0}$, the population weight $w_{i,j}$ is used to construct population-weighted neighbor variables.

Third and finally, utilizing relatively long pre-intervention periods in our sample, we also investigate the following model which further includes individual linear trends:

$$Y_{i,t} = \alpha_i + \beta_t + \tau Y_{-i,t} + \mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t + \mathbf{X}'_{-i,T_0} \cdot \boldsymbol{\pi}_t + \gamma_i \cdot t + \varepsilon_{i,t}, \tag{5}$$

where $\gamma_i \cdot t$ captures the individual linear trends that cannot be explained by observed covariates.

The roles of additional covariates and individual linear trends become clear when we consider two-step procedures to estimate τ. Because $Y_{-i,t}$ is endogenously determined in equations (3), (4) and (5) due to strategic interaction among municipalities (Brueckner 2003), we adopt a two stage least square (TSLS) approach where the first-stage and reduced-form models are described as follows when the model (5) is used:

$$\text{First stage: } Y_{-i,t} = \alpha_i^f + \beta_t^f + \theta Z_{-i,t} + \mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t^f + \mathbf{X}'_{-i,T_0} \cdot \boldsymbol{\pi}_t^f + \gamma_i^f \cdot t + \epsilon_{i,t}, \tag{6}$$

$$\text{Reduced form: } Y_{i,t} = \alpha_i^r + \beta_t^r + \phi Z_{-i,t} + \mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t^r + \mathbf{X}'_{-i,T_0} \cdot \boldsymbol{\pi}_t^r + \gamma_i^r \cdot t + \omega_{i,t}. \tag{7}$$

In these models, we use the same symbols for most coefficients as in equation (5), but with the indicators f and r at the upper right of the coefficients. $Z_{-i,t}$ is an instrumental variable defined as equation (1) and ϵ and ω are random errors.

The above two models are essentially difference-in-differences (DID) models where θ and ϕ capture the effect of neighbor amalgamation on the population-weighted average of insurance tax levels in neighbor and own municipalities respectively. In these models, pre-determined covariates $\mathbf{X}_{i,T_0}$ and $\mathbf{X}_{-i,T_0}$ as well as individual time trends are meant to control for differential trends that could violate the parallel trend assumption of DID.

Our parameter of interest τ in equation (5) can be obtained as a DID-IV estimate $\widehat{\tau} = \widehat{\phi}/\widehat{\theta}$. Here the reduced-form estimate $\widehat{\phi}$ can be interpreted as an estimate of

coefficients in a panel-data model, $\mathbf{X}'_{i,t} \cdot \boldsymbol{\mu} + \mathbf{X}'_{-i,t} \cdot \boldsymbol{\pi}$, because $\mathbf{X}_{i,t}$ and $\mathbf{X}_{-i,t}$ with $t = T_0+1$ or later are post-treatment variables that may be affected by our instrument of neighbor amalgamation and a primary motivation of the introduction of X_{i,T_0} and X_{-i,T_0} is to establish the conditional independence of our instrument in first-stage and reduced-form estimations, which is introduced in this subsection. A similar method of conditioning on covariates is adopted by, among others, Duflo (2004), who includes a vector of initial conditions with time-varying coefficients in her empirical specification.

the intention-to-treat (ITT) effect of neighbor amalgamation on own insurance tax levels. By weighting this estimated ITT effect with the first-stage estimate $\widehat{\theta}$ we can recover the effect of neighbor insurance tax levels on own insurance tax level, which we interpret as a parameter of tax mimicking.

Finally, in actual TSLS estimation we use the interactions of initial conditions and year dummies to estimate time-varying coefficients $\boldsymbol{\mu_t}$ and $\boldsymbol{\pi_t}$, using T_0 as a reference or initial year. Namely, the terms $\mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t$ and $\mathbf{X}'_{-i,T_0} \cdot \boldsymbol{\pi}_t$ in equation (3) are expressed as $\sum_{l \neq T_0} (\mathbf{X}_{i,T_0} \times Year_l)' \cdot \boldsymbol{\mu}_t$ and $\sum_{l \neq T_0} (\mathbf{X}_{-i,T_0} \times Year_l)' \cdot \boldsymbol{\pi}_t$, where $Year_l$ is a year dummy variable for year l.

4 Data

4.1 Data Construction

Our data construction procedures can be described as follows. First, we construct the panel data of municipalities (i.e. CHI insurers) that did not amalgamate during the sample period (1994-2007). This is done because otherwise municipalities would not be identical before and after the amalgamations and we would not be able to apply a DID method to them. Consequently, our data covers 1193 municipalities for 14 years.[9]

Second, we define neighbor municipalities based on municipal borders in 2002, just before the Great Heisei Amalgamation. Based on this definition of contiguity, we calculate a population weight of neighbor municipalities $w_{i,j}$ with equation (2), which is used for $Z_{-i,t}$, $Y_{-i,t}$ and X_{-i,T_0}. By adopting this procedure, we can construct $Y_{-i,t}$ so that it only reflects the changes in neighbor insurance tax levels, not the scale or number of neighbor amalgamations. [10]

[9]As for the municipalities that experienced amalgamations before the Great Heisei Amalgamation, we also exclude them from the sample but the number of such amalgamations is small: the number of municipalities decreased by only 26 during the period from April 1994 to April 2002. In addition, isolated islands are also excluded since the definition of "neighbor" municipalities with a common boundary cannot be straightforwardly applied to them. 34 municipalities are excluded based on this criteria.

[10]For instance, suppose that non-amalgamated municipality A has only one neighbor municipality B in 2002. In 2003 municipality B amalgamates with municipality C, which shares a border with B but not A. Thus a new municipality D, which consists of old municipalities B and C, is established in 2003. Consider the case in which the insurance tax level in the area of former municipality B does not change when municipality B became

Third, because the panel data of detailed insurance tax bases and individuals' incomes and assets is not available, we use the sum of the per-enrollee-based and per-household based elements of insurance tax for an outcome variable. We call this the "lump-sum insurance tax level", because the same tax burden is imposed on CHI enrollees if their household sizes are the same, regardless of their incomes and assets. Despite the incompleteness of this indicator as a proxy for the total insurance tax level for individual CHI enrollees, this variable is preferable for a cross-municipality comparison because of its simplicity and similarity to a simple lump-sum tax. In particular, lump-sum insurance tax levels may be highly correlated with the total insurance tax burden for the low-income households whose tax burdens from income-based and asset-based tax elements are relatively low.

Fourth, we use data from the Report on Citizens' Health Insurance (Kokumin kenko Hoken No Jittai) from 1994 to 2007, which are published by the All-Japan Federation of National Health Insurance Organizations. These data cannot be used after 2007 because insurance tax levels change discontinuously in 2008 due to the start of the Latter-stage Elderly Healthcare Program, which was a new program for the elderly aged 75 years and older introduced in 2008. The number of enrollees also changed discontinuously before and after this reform since people over 75 years of age in CHI were then enrolled in the Latter-stage Elderly Healthcare Program. We therefore restrict the sample period for our analysis from 1994 to 2007, which covers the entire period during which the Great Heisei Amalgamation was implemented.

Fifth and finally, 2002 is set as the pre-treatment reference year of covariates T_0.

4.2 Descriptive Statistics

Table 1 presents descriptive statistics of the variables that are used in our estimation. Note that our outcome variable, its spatially lagged variable, and the two instruments have a panel structure. As mentioned above, pre-determined covariates for own and neighbor municaplities are fixed to year 2002. When covariates are introduced in estimation models we take logs of all covariates except for adjustment transfer ratio, extra transfer ratio and discretionary transfer ratio, which sometimes

municipality D in 2003. In this case, when we use fixed population weights and the fixed definition of municipal proximity based on 2002, the variable $Y_{-i,t}$ does not change from 2002 to 2003. On the other hand, if we adopt contemporary populations and municipal proximity based on 2003 to construct $Y_{-i,t}$, $Y_{-i,t}$ changes from 2002 to 2003 even when neighbor insurance tax levels do not change during this period.

take a value of zero.

Table 1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Main outcome (1994-2007)					
Own lump-sum insurance tax level	16254	53620.87	13182.77	14880	123600
Outcomes used in Section 7 (1994-2007)					
Discretionary transfers per enrollee (total)	16254	1764.50	6965.52	0.00	464488.60
Discretionary transfers per enrollee (general account)	16254	884.94	4154.59	0.00	220466.00
Discretionary transfers per enrollee (CHI funds)	16254	879.56	5365.29	0.00	450929.10
Spatially lagged outcome (1994-2007)					
Neighbor lump-sum insurance tax level	16254	52847.34	11968.53	4628.415	87982.98
Instrumental variable (1994-2007)					
Intensity of neighbor amalgamation	16254	0.182	0.386	0	1
Own covariates (2002)					
Population	1161	57265.64	184237	545	3466875
CHI enrollee ratio	1161	40.272	8.574	18.472	77.911
Elderly ratio	1161	29.734	7.491	12.112	56.956
Retired ratio	1161	11.511	4.202	0.590	24.348
CHI revenue per capita	1161	26.705	48.180	0.057	586.315
Adjustment transfer ratio	1161	7.123	4.107	0.000	23.525
Extra transfer ratio	1161	1.108	2.942	0.000	44.859
Discretionary transfer ratio	1161	0.760	2.944	0.000	68.197
Growth rate of population (94-02)	1161	-1.569	9.303	-97.849	82.004
Growth rate of CHI enrollee ratio (94-02)	1161	14.891	11.346	-33.192	60.124
Growth rate of elderly ratio (94-02)	1161	5.438	3.111	-10.725	27.176
Growth rate of retired ratio (00-02)	1161	2.763	7.982	-38.323	59.184
Growth rate of CHI revenue per capita	1161	8.215	201.822	-27.478	4861.124
Population-weighted neighbor covariates (2002)					
Population	1161	139786	244919.3	1518.819	2059683
CHI enrollee ratio	1161	38.219	6.483	24.529	65.645
Elderly ratio	1161	29.020	5.739	12.458	49.528
Retired ratio	1161	12.112	3.053	2.453	22.632
CHI revenue per capita	1161	12.794	17.996	0.271	182.664
Adjustment transfer ratio	1161	6.940	3.542	0.146	22.578
Extra transfer ratio	1161	0.506	0.903	0.006	8.844
Non-statutory transfer ratio	1161	0.346	0.781	0.000	15.652

Notes: The sample consists of municipalities that did not amalgamate during the Great Heisei Amalgamation. The neighbor lump-sum insurance tax level is constructed as the population-weighted average of both amalgamated and non-amalgamated municipalities, where the population weight is calculated as $w_{i,j}$ in equation (2) with the reference year T_0 being set as 2002. The intensity of neighbor amalgamation is based on equation (1). Population-weighted neighbor covariates are also constructed as the population-weighted averages of both amalgamated and non-amalgamated municipalities with the population weight $w_{i,j}$.

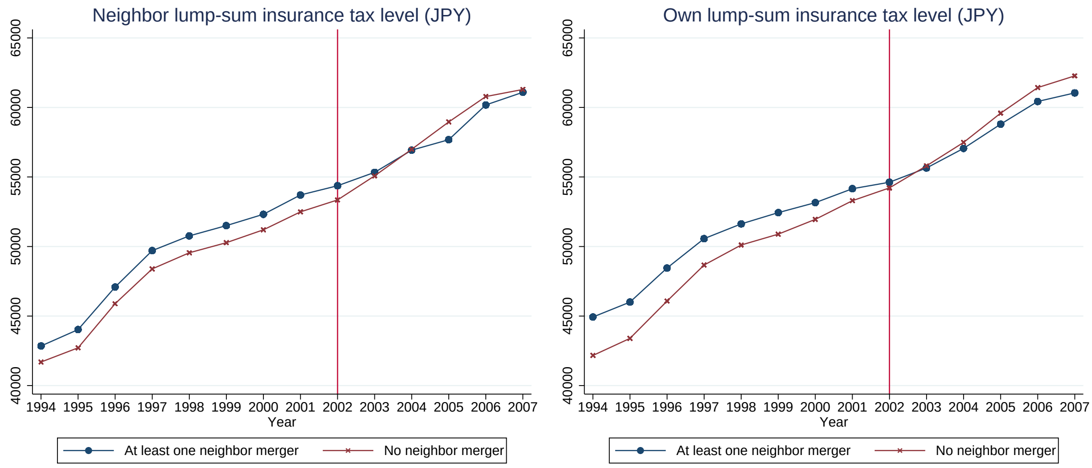
Figure 4 shows the descriptive trends of lump-sum insurance tax levels in neighbor (left-hand side) and own (right-hand side) municipalities for non-amalgamated municipalities, grouped by whether or not they were exposed to neighbor amalgamation by 2007. That is, municipalities with $Z_{-i,2007} = 0$ are one group (i.e. control) and municipalities with $Z_{-i,2007} > 0$ are the other group (i.e. treated)[11]. These graphs loosely correspond to the first-stage and reduced-form DID estimations respectively, so they can be utilized to confirm whether crude insurance tax level trends in neighbor and own municipalities satisfy the common trend assumption and whether the first-stage IV relevance and the reduced-form effect are observable in a graphical representation.

First, the left-hand graph shows the trends of neighbor lump-sum insurance tax levels for the municipalities that had had at least one neighbor amalgamation by 2007 ("treated") and the municipalities that had had no neighbor amalgamation by 2007 ("control"). The line at year 2002 indicates the beginning of the Great Heisei Amalgamation. This graph corresponds to the first-stage DID estimation with model (6), but with a simplified time-invariant binary instrument. This graph shows that the levels of lump-sum insurance tax differ between these two groups before 2002, implying that the incidence of neighbor amalgamation is not randomly assigned. On the other hand, the trends of the two tax levels appears to be similar before 2002, which suggests that the common trend assumption of the DID strategy seems valid without conditioning on additional covariates. In addition, although both trends exhibit upward trends after 2002, the upward trend of neighbor lump-sum insurance tax levels for "treated" municipalities seems to have been suppressed compared with that for "control" municipalities. This implies that our instrument of neighbor amalgamation satisfies the assumption of first-stage IV relevance.

Second, the right-hand graph presents the lump-sum insurance tax level trends for the same two groups as in the left-hand graph. This corresponds to the reduced-form DID estimation with model (7). At first glance the implications of this graph are similar to those of the left-hand graph. That is, the levels of lump-sum insurance tax clearly differ between the two groups in the pre-treatment period, but their trends look more or less similar in the period examined. At the same time, the upward trend in insurance tax levels for "treated" municipalities appears to be less steep than that for the "control" municipalities even before 2002, when the Great Heisei Amalgamation started. Thus a straightforward DID estimation may not produce a reliable causal estimate without controlling for this differential trend. As discussed in Section 3.4, we incorporated several own covariates, population-

[11]Note that the sample average of $Z_{-i,2007}$ is 0.368 (S.D. 0.362) and the percentage of municipalities with $Z_{-i,2007} > 0$ is about 70.3%.

Figure 4: Trends of lump-sum insurance tax levels



Notes: Only municipalities that did not amalgamate between 1994-2007 are included in the sample. Insurance tax levels are averaged in each group. The left-hand graph corresponds to "first-stage DID" and the right-hand graph corresponds to "reduced-form DID" with a simplified binary instrument of neighbor amalgamation.

weighted neighbor covariates, and individual linear trends into the regressors. See also Section 7, where we present how our DID strategy works with placebo tests.

5 Results

5.1 Baseline results

Results of TSLS estimation using the whole sample are shown in the panel (A) in Table 2. The first, second, and third columns show estimation results with the models (3), (4) and (5) respectively. To begin with, the first-stage estimates are statistically significant at the 1% significance level and the first-stage F statistics are around 10 or more in all three models. The first-stage estimates show that a 100 percent increase in the share of neighbor amalgamation (i.e. from no neighbor amalgamation to full neighbor amalgamation) leads to a reduction in population-averaged neighbor insurance taxes of around 2,000 JPY (about 24 USD) per year.

When it comes to TSLS estimates, they are around one and significantly different from zero in columns (I) and (II). In column (III), however, the estimate is about 0.3 and not statistically significant. This implies that the models (3) and (4) may not successfully mitigate the estimation bias that is generated by the non-random

Table 2: Baseline TSLS estimates

	(I)	(II)	(III)
(A) Sample: all observations			
TSLS estimate	0.973**	1.038**	0.296
	(0.389)	(0.504)	(0.270)
Observations	16,254	16,254	16,254
N of municipalities	1,161	1,161	1,161
First-stage estimate	-1,788.542***	-1,445.530***	-1,918.109***
	(501.821)	(496.402)	(460.722)
First-stage F stat.	12.70	8.480	17.33
(B) Sample: higher tax levels than neighbors in 2002			
TSLS estimate	2.715	3.289	0.996**
	(1.868)	(2.584)	(0.454)
Observations	8,414	8,414	8,414
N of municipalities	601	601	601
First-stage estimate	-992.592	-824.345	-1,735.906***
	(693.732)	(666.674)	(648.234)
First-stage F stat.	2.047	1.529	7.171
(C) Sample: lower tax levels than neighbors in 2002			
TSLS estimate	0.240	0.071	-0.254
	(0.361)	(0.440)	(0.400)
Observations	7,840	7,840	7,840
N of municipalities	560	560	560
First-stage estimate	-2,641.986***	-2,086.786***	-1,921.318***
	(711.205)	(722.448)	(618.530)
First-stage F stat.	13.80	8.343	9.649
Covariates	No	Yes	Yes
Individual trends	No	No	Yes

Note: ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. A robust standard error clustered by municipality is presented in parentheses.

occurrence of neighbor amalgamation, whereas individual linear trends incorporated into the model (5) may eliminate this bias and result in a statistically insignificant estimate. Figure 4 in the last section partly supports this view because the upward trends in the pre-treatment period look steeper for municipalities with no neighbor amalgamation. If we do not control for this trend differential a reduced-form

DID estimate should have an upward bias. Our preferred model is therefore equation (5) that incorporates individual linear trends to mitigate pre-treatment trend differentials between municipalities.

In panels (B) and (C) in Table 2 we provide TSLS estimates with the same model specifications as before but using the subsamples of municipalities whose lump-sum insurance tax levels were higher or lower than those of neighbor municipalities in 2002. The reason we use these subsamples is that we expect that tax mimicking behavior may be strong in municipalities whose insurance tax levels are already higher than those in neighbor municipalities just before these neighbor municipalities reduce their insurance tax levels.

In panel (B), TSLS estimates using the subsample of municipalities in which lump-sum insurance tax levels were higher than or equal to neighbor levels in 2002 (hereafter "higher-tax" municipalities) are provided. In columns (I) and (II), TSLS estimates are unreasonably high, possibly due to the weak IV problem (i.e. these first-stage estimates are not statistically significant), and statistically not different from zero. On the other hand, a TSLS estimate using our preferred model (5) in column (III) is around 1.0 and statistically significant. The first-stage estimate is statistically significant and the first-stage statistic is not very low (around seven), implying that weak instrument bias, if it exists, should not be a serious problem here.

Panel (C) shows TSLS estimates with the subsample of municipalities in which lump-sum insurance tax levels were lower than neighbor levels in 2002 (hereafter "lower-tax" municipalities). Because the municipalities in this subsample already had lower insurance tax levels than their neighbors before the Great Heisei Amalgamation, we expect no tax mimicking for these municipalities even when neighbor municipalities suppress insurance tax levels. The estimation results show that first-stage F statistics are around 10 and the TSLS estimates are all statistically not different from zero regardless of model specifications.

The estimation results with the model (5) shown in column (III) support the hypothesis of tax mimicking only for the subgroup of municipalities which had higher tax levels than those of neighbor municipalities. That is, municipalities that had higher tax levels than their neighbors respond to their neighbors' tax reduction, but municipalities that already had lower tax levels than their neighbors do not. Overall, if we believe that the model (5) is plausible for our subsample analysis, this surprisingly sharp difference in the TSLS estimates between panels (B) and

(C) implies quite heterogeneous tax mimicking behavior among municipalities.

5.2 Robustness checks

One interesting finding in our baseline analysis with Table 2 is the distinct contrast between the almost one-for-one effect in the "higher-tax" municipalities (panel (B)) and the zero effect for "lower-tax" municipalities (panel (C)). In order to check whether this finding is robustly observed, we estimate the coefficient of tax mimicking with different model specifications and subsamples.

The first robustness check relies on the supposition that the introduction of individual linear trends may eliminate most endogeneity bias in the first-stage and the TSLS estimation in panel (B) of Table 2, because the first-stage and TSLS estimates significantly change once individual linear trends are incorporated into the model. If this is true, whether or not additional observed covariates are included may not change the first-stage and reduced-form estimates of interest once individual linear trends are controlled for.

Table 3 provides TSLS estimates using models that all incorporate individual linear trends but different sets of covariates. Panel (A) shows that TSLS estimates for "higher-tax" municipalities are robustly around one in all model specifications whereas panel (B) indicates that TSLS estimates for "lower-tax" municipalities are all not significantly different from zero. These results do not imply that our TSLS estimation does not suffer from omitted variables bias, but they provide indirect evidence that the TSLS estimates we obtained may be insensitive to other unobserved covariates once individual linear trends are controlled for.

In the second robustness check, we restrict the subsample of both "higher-tax" and "lower-tax" municipalities to more homogeneous groups with seemingly similar incentives for tax mimicking. That is, in a stepwise manner we exclude observations in which insurance tax levels are further away from neighbor insurance tax levels.

Results are shown in Table 4. Panel (A) provides TSLS estimates for the subsample of "higher-tax" municipalities in which observations are further limited by the condition that own insurance tax level is lower than or equal to neighbor insurance tax level $\times$ 1.5, 1,4, ..., or 1.1.[12] The TSLS estimates are robustly around one and mostly significantly different from zero regardless of sample selection. Panel

[12] That is, the subsample is restricted to the municipalities that satisfy the following condition: neighbor tax level $\leq$ own tax level $\leq$ neighbor tax level $\times$ 1.5, 1,4, ..., or 1.1.

Table 3: TSLS estimates using different covariates

	(I)	(II)	(III)	(IV)	(V)
	(A) Sample: higher tax levels than neighbors in 2002				
TSLS estimate	0.836	1.018**	0.965**	1.128*	0.996**
	(0.553)	(0.516)	(0.448)	(0.615)	(0.454)
Observations	8,414	8,414	8,414	8,414	8,414
N of municipalities	601	601	601	601	601
First-stage estimate	-1,337.138**	-1,603.691**	-1,724.190***	-1,461.521**	-1,735.906***
	(669.093)	(688.717)	(652.564)	(709.127)	(648.234)
First-stage F stat.	3.994	5.422	6.981	4.248	7.171
	(B) Sample: lower tax levels than neighbors in 2002				
TSLS estimate	-0.634	-0.291	-0.348	-0.218	-0.254
	(0.545)	(0.461)	(0.475)	(0.400)	(0.400)
Observations	7,840	7,840	7,840	7,840	7,840
N of municipalities	560	560	560	560	560
First-stage estimate	-1,596.591***	-1,667.289***	-1,626.893***	-1,919.978***	-1,921.318***
	(596.221)	(628.908)	(623.257)	(617.252)	(618.530)
First-stage F stat.	7.171	7.028	6.814	9.675	9.649
Covariates					
Own	No	Yes	Yes	Yes	Yes
Own (growth rate)	No	No	Yes	No	Yes
Neighbor	No	No	No	Yes	Yes
Individual trends	Yes	Yes	Yes	Yes	Yes

Note: ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. A robust standard error clustered by municipality is presented in parentheses.

(B) in turn presents estimates for the subsample of "lower-tax" municipalities in which a counterpart lower limit is used for further selection of observations [13] In this case, The TSLS estimates are robustly not significantly different from zero. These results also support our baseline estimation results.

[13] That is, the following condition is used for subsample selection: neighbor tax level $\times$ 0.5, 0.6, ..., or 0.9 $\leq$ own tax level $<$ neighbor tax level.

Table 4: TSLS estimates with different subsamples

	(I)	(II)	(III)	(IV)	(V)
	(A) Sample: higher tax levels than neighbors in 2002				
TSLS estimate	0.959**	1.009**	1.245**	0.881**	0.810
	(0.443)	(0.459)	(0.559)	(0.359)	(0.530)
Observations	8,288	8,218	7,756	6,678	4,046
N of municipalities	592	587	554	477	289
First-stage estimate	-1,743.049***	-1,725.391***	-1,622.294**	-2,190.014***	-1,838.758**
	(649.943)	(651.988)	(667.729)	(687.704)	(927.895)
First-stage F stat.	7.192	7.003	5.903	10.14	3.927
Sample, upper limit	≤ Nei. tax ×1.5	≤ Nei. tax ×1.4	≤ Nei. tax ×1.3	≤ Nei. tax ×1.2	≤ Nei. tax ×1.1
	(B) Sample: lower tax levels than neighbors in 2002				
TSLS estimate	-0.254	-0.228	-0.154	-0.185	-0.464
	(0.400)	(0.402)	(0.444)	(0.413)	(0.583)
Observations	7,840	7,798	7,420	6,762	4,242
N of municipalities	560	557	530	483	303
First-stage estimate	-1,921.318***	-1,903.548***	-1,654.997**	-1,823.617***	-1,672.760**
	(618.530)	(621.097)	(645.333)	(679.259)	(792.998)
First-stage F stat.	9.649	9.393	6.577	7.208	4.450
Sample, lower limit	≥ Nei. tax ×0.5	≥ Nei. tax ×0.6	≥ Nei. tax ×0.7	≥ Nei. tax ×0.8	≥ Nei. tax ×0.9

Note: ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. A robust standard error clustered by municipality is presented in parentheses.

6 Placebo analysis

6.1 Two threats to the research design

In the last section, we interpret the estimation results based on our preferred model (5) as causal effects of strategic municipal interaction. There are, however, at least two major threats to the internal validity of our identification strategy.

First, as in the other recent quasi-experimental studies of mimicking behavior, our identifying variation of neighbor amalgamation may not be exogenous, and controlling for observed covariates, unobserved fixed effects, and individual linear trends may not sufficiently alleviate this problem.

In the context of the identification of spatial or strategic interactions between local governments, most identifying variations used in the previous studies seem to have some spatial correlation. This is because external reforms or shocks, which are often used to facilitate quasi-experiments, likely reflect the characteristics of local governments, and in turn the characteristics of local governments are often spatially correlated.

Thus, although it seems plausible to argue that spatially lagged covariates are not valid instruments because they may be correlated with the outcome variable through other passes than a spatially lagged outcome (Gibbons and Overman 2012), this criticism also, to some degree, applies to many quasi-experimental identifying variations. We then need to explicitly discuss (1) to what degree an exploited source of quasi-experimental variation solves or alleviates the problem of endogeneity bias in the estimation of strategic interaction and (2) how we can further reduce bias when some quasi-experimental variation is exploited for identification.

In our case, the instrument of neighbor amalgamation may also be endogenous. Namely, although the incidence of neighbor amalgamation appears to occur somewhat exogenously from the perspective of a non-amalgamated municipality, the intensity of neighbor amalgamation is presumably determined by the various socioeconomic and geographic circumstances of surrounding municipalities and possibly those of the municipality in question itself. This "endogenous neighbor amalgamation" may then result in estimation bias.

Second, using the subsample of non-amalgamates may also be problematic because this sample selection procedure is based on post-treatment status. That is, our sample consists of the municipalities that "survive" the Great Heisei Amalgamation with the municipalities that amalgamated during the period having been dropped. In fact, because municipalities in our sample are defined based on municipal borders in 2002, municipalities that amalgamated after 2002 cannot be well defined in our sample. This may cause a problem called "censoring by death" in the causal inference literature (Frangakis and Rubin 2002; Rubin 2006). In our case this problem can be termed "censoring by amalgamation," and this censoring can lead to estimation bias even when the problem of endogenous neighbor amalgamation does not exist.[14]

[14]The reason that censoring by amalgamation is problematic can be intuitively explained in a simplified setting. Suppose that the instrument is binary, that is, whether a non-amalgamated municipality in our sample is faced with neighbor amalgamation (treated group) or not (control group). Assume that municipalities consist of two response types: municipalities that would merge if their neighbor municipalities were willing to merge ("compliers") and municipalities that never merge regardless of neighbor municipalities' willingness to merge ("never-takers"). In this case, there are both compliers and never-takers in the control group. On the other hand, there are only never-takers in the treated group because the municipalities that experience their own (and therefore

Under the DID setting, both the problems of "endogenous neighbor amalgamation" and "censoring by amalgamation" may result in differential trends in neighbor insurance tax levels and own insurance tax levels among non-amalgamated municipalities with different intensities of neighbor amalgamation. In fact, the graphs in Figure 4 imply that a simple DID strategy may be sufficient to alleviate these problems for neighbor insurance tax levels (left-hand graph), but not for own insurance tax levels (right-hand graph); different trends seem to exist in the two groups before 2002 in the right-hand graph of own insurance tax levels.

6.2 Placebo tests

Assuming that the exclusion restriction of neighbor amalgamation is satisfied conditional on suitable pre-amalgamation confounding factors, the above threats can be alleviated by controlling for observed covariates, eliminating fixed effects, and imposing some parametric assumptions. In our case, we adopt a DID strategy as well as the parametric models expressed in equation (3)-(7) to eliminate possible bias caused by these threats.

Exploiting the DID model with time-varying treatment effects used in, among others, Autor (2003) and Finkeistein (2007), we can empirically show that there are no differential pre-treatment trends in neighbor and own insurance tax levels, controlling for observed covariates, municipality fixed effects, and municipality-specific linear trends.

To implement this placebo test, we need to modify the first-stage and reduced-form DID models (6) and (7) as follows:

$$\text{First stage: } Y_{-i,t} = \alpha_i^f + \beta_t^f + \theta_t Z_{-i,2007} + \mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t^f + \mathbf{X}'_{-i,T_0} \cdot \boldsymbol{\pi}_t^f + \gamma_i^f \cdot t + \epsilon_{i,t}, \quad (8)$$

$$\text{Reduced form: } Y_{i,t} = \alpha_i^r + \beta_t^r + \phi_t Z_{-i,2007} + \mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t^r + \mathbf{X}'_{-i,T_0} \cdot \boldsymbol{\pi}_t^r + \gamma_i^r \cdot t + \omega_{i,t}. \quad (9)$$

That is, we replace the terms $\theta Z_{-i,t}$ and $\phi Z_{-i,t}$ with $\theta_t Z_{-i,2007}$ and $\phi_t Z_{-i,2007}$, where θ_t and ϕ_t are time-varying DID coefficients and $Z_{-i,2007}$ is a time-invariant instru-

neighbor) amalgamation are dropped from our sample. Then the difference in insurance tax levels between the treated and control groups cannot be interpreted as a reduced-form effect of neighbor amalgamation on tax levels because the treated group and the control group consist of different types of municipalities. A similar problem would presumably still exist in a more realistic setting with a continuous indicator of neighbor amalgamation and municipalities with more complicated types of responses, and might result in estimation bias.

ment that is an indicator of neighbor amalgamation intensity in 2007.[15]

Coefficients θ_t and ϕ_t are meant to capture the varying effects of $Z_{-i,2007}$ on neighbor and own lump-sum insurance tax levels respectively. We expect that both the estimates of θ_t and ϕ_t will be around zero before 2002 and then gradually become negative if differential trends are successfully controlled for, the instrument is relevant, and tax mimicking exists. In actual estimation models, $\sum_{l \neq T_0} \theta_t \cdot (Z_{-i,2007} \times Year_l)$ and $\sum_{l \neq T_0} \phi_t \cdot (Z_{-i,2007} \times Year_l)$ are used for $\theta_t \cdot Z_{-i,2007}$ and $\phi_t \cdot Z_{-i,2007}$ to obtain time-varying DID estimates, where the reference year T_0 is set as 2002.

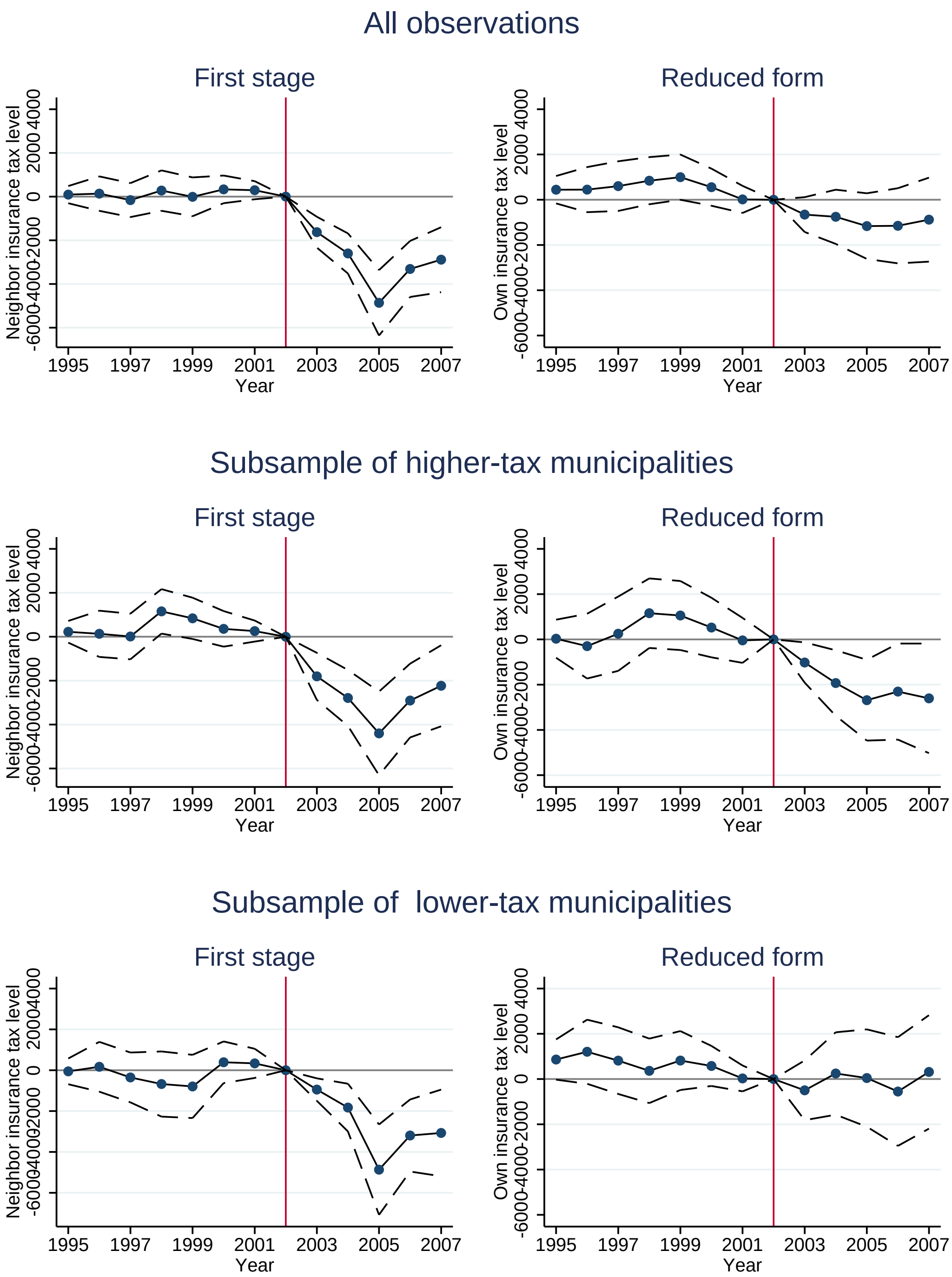
Figure 5 provides the results of the placebo tests, using the whole sample, the subsample of "higher-tax" municipalities, and the subsample of "lower-tax" municipalities. Our primary interest is whether estimates are around zero before the period of Great Heisei Amalgamation in both the first-stage and reduced-form DID analysis. Estimates and their confidence intervals in all six graphs show that both first-stage and reduced-form estimates are not significantly different from zero in the pre-amalgamation period (before 2002), implying that bias-inducing differential trends during this period do not exist or have been effectively controlled for.

In addition, first-stage estimates during the amalgamation period (2003-2007) in all three cases suggest that neighbor amalgamation starts suppressing the average tax levels of adjacent municipalities in 2003, implying that our instrument's first-stage assumption is quite plausible. Reduced-form estimates during the same period are also informative because they show that neighbor amalgamation only results in the suppression of lump-sum insurance tax levels for "higher-tax" municipalities. For the subsample of "lower-tax" municipalities, we do not find any significant changes in estimates before and after 2002. This result is expected from Table 2 and supports the hypothesis of tax mimicking behavior only occurring in "higher-tax" municipalities.

In Appendix A, we also provide the counterpart figures of Figure 5 based on TSLS models without covariates and individual trends (Figure (A.1)) and based on TSLS models with covariates but without individual trends (Figure (A.2)). They show that first-stage placebo estimates (i.e. estimates before 2002) are around zero and show no differential trends in the pre-amalgamation period without controlling for individual trends, but the counterpart reduced-form placebo estimates exhibit some downward trends. This implies that controlling for individual linear trends is necessary in order for our identification strategy to obtain unbiased estimates.

[15]See footnote 11 for the mean and the standard deviation of $Z_{-i,2007}$.

Figure 5: DID estimates before and after Great Heisei Amalgamation



Note: These estimation results are based on TSLS models (8) and (9). Solid lines with dots present point estimates and dashed lines provide 95% confidence intervals. The confidence intervals are calculated by robust standard errors clustered by municipality.

7 Further analysis

7.1 Mechanism

If tax mimicking behavior exists, how are municipalities able to manipulate their insurance tax levels? As we have already argued, one plausible way to reduce insurance tax levels is to utilize discretionary transfers from other fiscal resources. To investigate whether the amount of discretionary transfers is affected by neighbor amalgamation is therefore a good alternative method of checking for tax mimicking behavior.

We reanalyze the empirical model (5) using the alternative outcome variable of discretionary transfers $G_{i,t}$. Then the reduced-form model (7) is replaced with the following model:

$$G_{i,t} = \alpha_i^r + \beta_t^r + \phi Z_{-i,t} + \mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t^r + \mathbf{X}'_{-i,T_0} \cdot \boldsymbol{\pi}_t^r + \gamma_i^r \cdot t + \omega_{i,t}, \tag{10}$$

where the first-stage model is identical to model (6). An intuition behind this TSLS estimation is that there should also be an effect of neighbor amalgamation on discretionary transfers if the effect of neighbor insurance tax levels on own insurance tax levels is mediated through discretionary transfers to CHI accounts.

Results are shown in Table 5. As Table 2 indicates, we provide TLS estimates with the whole sample (A), the subsample of "higher tax" municipalities (B), and the subsample of "lower tax" municipalities (C). The first, second, and third columns show the estimation results for different fiscal variables, that is, the total amount of discretionary transfers, discretionary transfers from municipal general accounts, and discretionary transfers from CHI funds.[16] Note that the magnitudes of estimates reported as our main results in Section 5 and those reported in this subsection cannot be directly compared (the former indicate effects on institutional lump-sum tax levels whereas the latter represent effects on per-enrollee revenue values).

Panel (A) shows that neighbor insurance tax levels negatively affect discretionary transfers and this effect is found exclusively in transfers from CHI funds. That is, the decrease in neighbor insurance tax levels may lead to an increase in discretionary transfers from CHI funds, but not from municipal general accounts.

Similar results are found in panel (B), where the subsample of "higher-tax"

[16] See Table 1 for the descriptive statistics of these outcome variables.

Table 5: TSLS estimates for discretionary transfer

	(I)	(II)	(III)
	Y: discretionary transfers		
Sample	Total	From municipal general accounts	From CHI funds
(A) Sample: all observations			
TSLS estimate	-0.462**	-0.058	-0.404**
	(0.213)	(0.129)	(0.198)
Observations		16,254	
N of municipalities		1,161	
First stage estimate		-1,910.081*** (460.722)	
First stage F stat.		17.33	
(B) Sample: higher tax levels than neighbors in 2002			
TSLS estimate	-0.365*	-0.018	-0.347
	(0.216)	(0.131)	(0.224)
Observations		8,414	
N of municipalities		601	
First stage estimate		-1,735.906***(648.234)	
First stage F stat.		7.171	
(C) Sample: lower tax levels than neighbors in 2002			
TSLS estimate	-0.553	-0.152	-0.401
	(0.336)	(0.218)	(0.282)
Observations		7,840	
N of municipalities		560	
First stage estimate		-1,952.383*** (618.530)	
First stage F stat.		9.649	

Note: ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. Total discretionary transfers for each municipality are the sum of discretionary transfers from a municipal general account and a CHI fund. All estimates are generated using a model that incorporates both additional covariates and individual linear trends. A robust standard error clustered by municipality is presented in parentheses.

municipalities is used, but the statistical significance is weaker and the magnitude of estimates is smaller than those in panel (A). In panel (C) the subsample of "lower-tax" municipalities are all statistically not different from zero at the 10 % significance level, but the estimates are more or less similar to those in panels (A) and (B).

We interpret these findings as follows. First, statistically significant effects are observed only on discretionary transfers from CHI funds because our analysis examines only the short-run impact of the suppression of neighbor tax levels during the Great Heisei Amalgamation. A municipality being faced with a sharp change in neighbor tax levels due to neighbor amalgamation may deal with it by using a reversal from its own CHI fund.[17] Second, both "higher-tax" and "lower-tax" municipalities allocate more discretionary transfers from CHI funds to CHI accounts when they are faced with a decrease in neighbor insurance tax levels caused by neighbor amalgamations, but their statistical power is too weak to robustly capture this effect. Third, only "higher-tax" municipalities utilize these discretionary transfers to suppress lump-sum insurance tax levels. "Lower-tax" municipalities may use the transfers for other purposes but not for the reduction of the lump-sum insurance tax.

When it comes to the third point, this speculation is plausible given the fact that "lower-tax" municipalities already have lower lump-sum insurance tax levels than neighbor municipalities when they are faced with the reduction of neighbor insurance tax levels. In this case, "lower-tax" municipalities nonetheless may increase discretionary transfers if other dimensions of the CHI tax change in neighbor municipalities as a result of amalgamation. For example, if neighbor amalgamation leads to the expansion of the special insurance exemption for low-income households or asset-based insurance tax levels in neighbor municipalities, "lower-tax" municipalities may use discretionary transfers to "mimic" such changes. Unfortunately, we cannot investigate these behaviors because we do not have panel data for any dimensions of the CHI tax other than its lump-sum elements.

7.2 Comparisons with SIV estimates

Recent quasi-experimental studies such as Lyytikäinen (2012), Baskaran (2014), Isen (2014), and Parchet (2014) compare estimates obtained from quasi-experimental research designs with estimates based on conventional spatial econometric approaches. In this section, we also provide counterpart estimates using a spatial instrumental variables (SIV) approach where spatially lagged covariates are used as instruments.

The first-stage and second-stage models for the SIV approach are expressed as follows:

$$\text{First stage: } Y_{-i,t} = \alpha_i^f + \beta_t^f + \mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t^f + \mathbf{X}'_{-i,T_0} \cdot \boldsymbol{\pi}_t^f + \epsilon_{i,t}, \tag{11}$$

[17] Transfers from CHI funds are meant to stabilize CHI accounts and may be more flexibly used for adjusting insurance tax levels in the short run than transfers from municipal general accounts.

$$\text{Second stage: } Y_{i,t} = \alpha_i^s + \beta_t^s + \tau Y_{-i,t} + \mathbf{X}'_{i,T_0} \cdot \boldsymbol{\mu}_t^s + \varepsilon_{i,t}, \tag{12}$$

where excluded instruments in (11) are the whole set of spatially lagged covariates $\mathbf{X}_{-i,T_0}$. To facilitate comparison to our main results in Section 5, we continue to use the same sample of non-amalgamates as in the previous sections. We also use the same time-invariant own and neighbor covariates $\mathbf{X}_{i,T_0}$ and $\mathbf{X}_{-i,T_0}$ as before, where T_0 is set as 2002. We can interpret these variables as exogenous time-fixed municipal factors with time-varying effects on lump-sum insurance tax levels. In contrast to our quasi-experimental TSLS models (6) and (7), the above SIV models do not include individual linear trend terms $\gamma_i^f \cdot t$ and $\gamma_i^s \cdot t$ because they are included in the DID models to strengthen the common trend assumption, which is not addressed in the current SIV setting. We do, however, also check SIV estimates with models that incorporate individual linear trends.

As in the baseline analysis, we also conduct subsample analysis with the municipalities that had higher or lower tax levels than their neighbors. Unlike the DID strategy adopted in the main estimation, we have no rationale for splitting the sample based on the levels of outcome variables in 2002. Under the SIV framework with the models (11) and (12), we exploit "within" variations in $Y_{-i,t}$ and $Y_{i,t}$ from the first year of the sample. We therefore divide the sample based on the levels of outcome variables in the first year (1994).

The results are presented in Table 6. Panel (A) shows TSLS estimates using the whole sample. The first column, an estimate obtained using a model without covariates $\mathbf{X}_{i,T_0}$ and individual trends, shows a statistically significant estimate of around 0.9. The estimate in the second column, in which we add covariates to the models, is also significantly different from zero (about 0.35). In the third column, in which we further incorporate individual linear trends into the models, however, the estimate is not significantly different from zero. First-stage F statistics are relatively low, especially in the second and third column, so we must interpret the results with caution. In panels (B) and (C) we provide TSLS estimates using subsamples and overall results that are more or less similar to the results of panel (A).

One primary finding from these results is that the SIV estimate is large and significantly different from zero when we do not control for individual linear trends. However this positive estimate completely disappears when individual linear trends are incorporated into the model, although this result should also be interpreted

Table 6: TSLS estimates with SIV

	(I)	(II)	(III)
(A) Sample: all observations			
TSLS estimate	0.875***	0.347***	0.109
	(0.072)	(0.085)	(0.113)
Observations	16,254	16,254	16,254
N of municipalities	1,161	1,161	1,161
First-stage F stat.	10.17	5.554	3.546
(B) Sample: higher tax levels than neighbors in 1994			
TSLS estimate	0.829***	0.438***	0.151
	(0.092)	(0.109)	(0.126)
Observations	8,932	8,932	8,932
N of municipalities	638	638	638
First-stage F stat.	7.476	3.744	2.236
(C) Sample: lower tax levels than neighbors in 1994			
TSLS estimate	0.876***	0.338***	0.049
	(0.093)	(0.110)	(0.155)
Observations	7,322	7,322	7,322
N of municipalities	523	523	523
First-stage F stat.	9.437	5.666	3.061
Covariates	No	Yes	Yes
Individual trends	No	No	Yes

Note: ***, **, and * denote significance at the 1%, 5%, and 10% level respectively. A robust standard error clustered by municipality is presented in parentheses.

with caution.[18]

To compare these SIV estimates with our benchmark quasi-experimental estimates in Table 2 is difficult. First, as discussed, the internal validity of the above SIV estimates seems to be limited. Second, even if we can assume that the SIV strategy, in particular its exclusion restriction assumption, is plausible, it is still

[18]If individual trends in $Y_{-i,t}$ and $Y_{i,t}$ across the sample period reflect tax mimicking among contiguous municipalities, the terms of individual linear trends can mistakenly eliminate this kind of tax mimicking from the variation in $Y_{-i,t}$ and $Y_{i,t}$. In contrast, the introduction of individual linear trends in our DID analysis should not be seriously affected by this problem because the terms of individual linear trends in DID analysis are meant to control for differential trends in the pre-amalgamation period among municipalities with different treatment intensity.

hard to understand what is being captured by these SIV estimates.

Assuming that SIV estimation is internally valid, one possible causal interpretation is that SIV estimates are capturing the tax mimicking behavior that is generated by exogenous within-municipality variation of spatially lagged covariates. If this is the case, the difference between a quasi-experimental estimate and a SIV estimate simply reflects the fact that both estimates represent different kinds of tax mimicking behavior: the former captures short-term responses to sudden reductions in neighbor tax levels caused by amalgamation, and the latter represents more general responses to changes in neighbor tax levels caused by changes in neighbor socio-economic factors.

Overall, although we can argue that a SIV approach is less plausible than a quasi-experimental approach as Gibbons and Overman (2012) advocate, the estimation results of our quasi-experimental study do not provide general implications about the existence and magnitude of strategic municipal interactions that the SIV approach tries to identify.

8 Conclusion

This paper utilizes the sharp reduction in Japanese municipal health insurance taxes during the Great Heisei Amalgamation in order to break down the simultaneous determinants of insurance tax levels among neighboring municipalities. Exploiting this quasi-experimental shock on neighbor insurance tax levels, our TSLS estimates find that a one point reduction in the weighted average of lump-sum insurance tax levels led to almost a one-for-one reduction in lump-sum insurance tax levels for the subgroup of municipalities that had higher lump-sum insurance tax levels than neighbor municipalities before the Great Heisei Amalgamation. We do not find tax mimicking behavior in the other municipalities.

These results are contrary to the findings of recent quasi-experimental studies in tax mimicking such as Lyytikäinen (2012), Baskaran (2014), Isen (2014), and Parchet (2014), which find no or only modest tax mimicking behavior. We argue that this difference is primarily due to the very limited external validity of these previous studies. As discussed in the Introduction, these previous studies often focus on situations in which tax mimicking is unlikely to occur. Quite strong short-run tax mimicking behavior observed in this study implies that short-run tax mimicking occurs in some specific situations, and whether or not a quasi-experimental study can find tax mimicking depends considerably on its research design and identifying variation.

Note that the lack of external validity is also one of the limitations of our study. Our sample consists of non-amalgamated municipalities during the period of the Great Heisei Amalgamation, and the excluded amalgamated municipalities amount to 30 percent of the total. Since the characteristics of non-amalgamates are different from those of amalgamates, our findings cannot be simply applied to other municipalities in Japan. In addition, the identifying variation of neighbor municipalities itself is in some sense a very specific case and we need to be careful to interpret our results in a more general context.

We would nevertheless stress that these drawbacks are not limited to our study, but applicable to most recent reliable quasi-experimental studies on tax mimicking. In addition, our estimation results indicate that local governments strongly respond to their neighbors when they seem to have incentives to do so. To clarify under what conditions local governments mimic their neighbors' tax levels and why they do so is an important task to be taken up in future research.

Bibliography

Agrawal, David R. (2015) "The Tax Gradient: Spatial Aspects of Fiscal Competition," *American Economic Journal: Economic Policy*, Vol. 7, No. 2, pp. 1–29.

Allers, Maarten a. and J. Paul Elhorst (2005) "Tax Mimicking and Yardstick Competition among Local Governments in the Netherlands," *International Tax and Public Finance*, Vol. 12, No. 1993, pp. 493–513.

Autor, David H. (2003) "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing," *Journal of Labor Economics*, Vol. 21, No. 1, pp. 1–42.

Baskaran, Thushyanthan (2014) "Identifying Local Tax Mimicking with Administrative Borders and a Policy Reform," *Journal of Public Economics*, Vol. 118, pp. 41–51.

Besley, Timothy and Anne Case (1995) "Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition," *The American Economic Review*, Vol. 85, No. 1, pp. 25–45.

Brueckner, Jan K. (2003) "Strategic Interaction Among Governments: An Overview of Empirical Studies," *International Regional Science Review*, Vol. 26, No. 2, pp. 175–188.

Brülhart, Marius and Raphaël Parchet (2014) "Alleged Tax Competition: The Mysterious Death of Bequest Taxes in Switzerland," *Journal of Public Economics*, Vol. 111, pp. 63–78.

Costa-Font, Joan, Filipe De-Albuquerque, and Hristos Doucouliagos (2014) "Do Jurisdictions Compete on Taxes? A Meta-regression Analysis," *Public Choice*, Vol. 161, No. 3-4, pp. 451–470.

Deaton, Angus (2010) "Instruments, Randomization, and Learning about Development," Vol. 48, No. June, pp. 424–455.

Duflo, Esther (2004) "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia," *Journal of Development Economics*, Vol. 74, No. 1, pp. 163–197.

Eugster, Beatrix and Raphaël Parchet (2014) "Culture and Taxes: Towards Identifying Tax Competition."

Finkeistein, Amy (2007) "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare," *The Quarterly Journal of Economics*, Vol. CXXII, No. August, pp. 1–37.

Frangakis, Constantine E and Donald B Rubin (2002) "Principal Stratification in Causal Inference.," *Biometrics*, Vol. 58, No. 1, pp. 21–29.

Gibbons, Stephen and Henry G. Overman (2012) "Mostly Pointless Spatial Econometrics?," *Journal of Regional Science*, Vol. 52, No. 2, pp. 172–191.

Heckman, James J (2000) "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective," *The Quarterly Journal of Economics*, Vol. 115, No. 1, pp. 45–97.

Hinnerich, Björn Tyrefors (2009) "Do Merging Local Governments Free Ride on Their Counterparts When Facing Boundary Reform?" *Journal of Public Economics*, Vol. 93, No. 5, pp. 721–728.

Ikegami, Naoki, Byung-Kwang Yoo, Hideki Hashimoto, Masatoshi Matsumoto, Hiroya Ogata, Akira Babazono, Ryo Watanabe, Kenji Shibuya, Bong-Min Yang, Michael R Reich, and Yasuki Kobayashi (2011) "Japanese Universal Health Coverage: Evolution, Achievements, and Challenges," *The Lancet*, Vol. 378, No. 9796, pp. 1106–1115.

Isen, Adam (2014) "Do Local Government Fiscal Spillovers Exist? Evidence from Counties, Municipalities, and School Districts," *Journal of Public Economics*, Vol. 110, pp. 57–73.

Jordahl, Henrik and Che-Yuan Liang (2010) "Merged Municipalities, Higher Debt: On Free-riding and the Common Pool Problem in Politics," *Public Choice*, Vol. 143, No. 1-2, pp. 157–172.

Lyytikäinen, Teemu (2012) "Tax Competition among Local Governments: Evidence from a Property Tax Reform in Finland," *Journal of Public Economics*, Vol. 96, No. 7-8, pp. 584–595.

Nakazawa, Katsuyoshi (2015) "Amalgamation, Free-rider Behavior, and Regulation," *International Tax and Public Finance*.

Nordhaus, William D (1975) "The Political Business Cycle," *The review of economic studies*, pp. 169–190.

Parchet, Raphaël (2014) "Are Local Tax Rates Strategic Complements or Strategic Substitutes?"Technical report, Università della Svizzera italiana.

Reingewertz, Yaniv (2012) "Do Municipal Amalgamations Work? Evidence from Municipalities in Israel," *Journal of Urban Economics*, Vol. 72, No. 2-3, pp. 240–251.

Rogoff, Kenneth (1990) "Equilibrium Political Budget Cycles."

Rubin, D B (2006) "Causal Inference through Potential Outcomes and Principal Stratification: Application to Studies with "Censoring" Due to Death," *Statistical Science*, Vol. 21, No. 3, pp. 299–309.

Takaku, Reo, Shunichiro Bessho, Shuzo Nishimura, and Naoki Ikegami (2014) "Fiscal Disparities among Social Health Insurance Programs in Japan," *Universal Health Coverage for Inclusive and Sustainable Development*, p. 41.

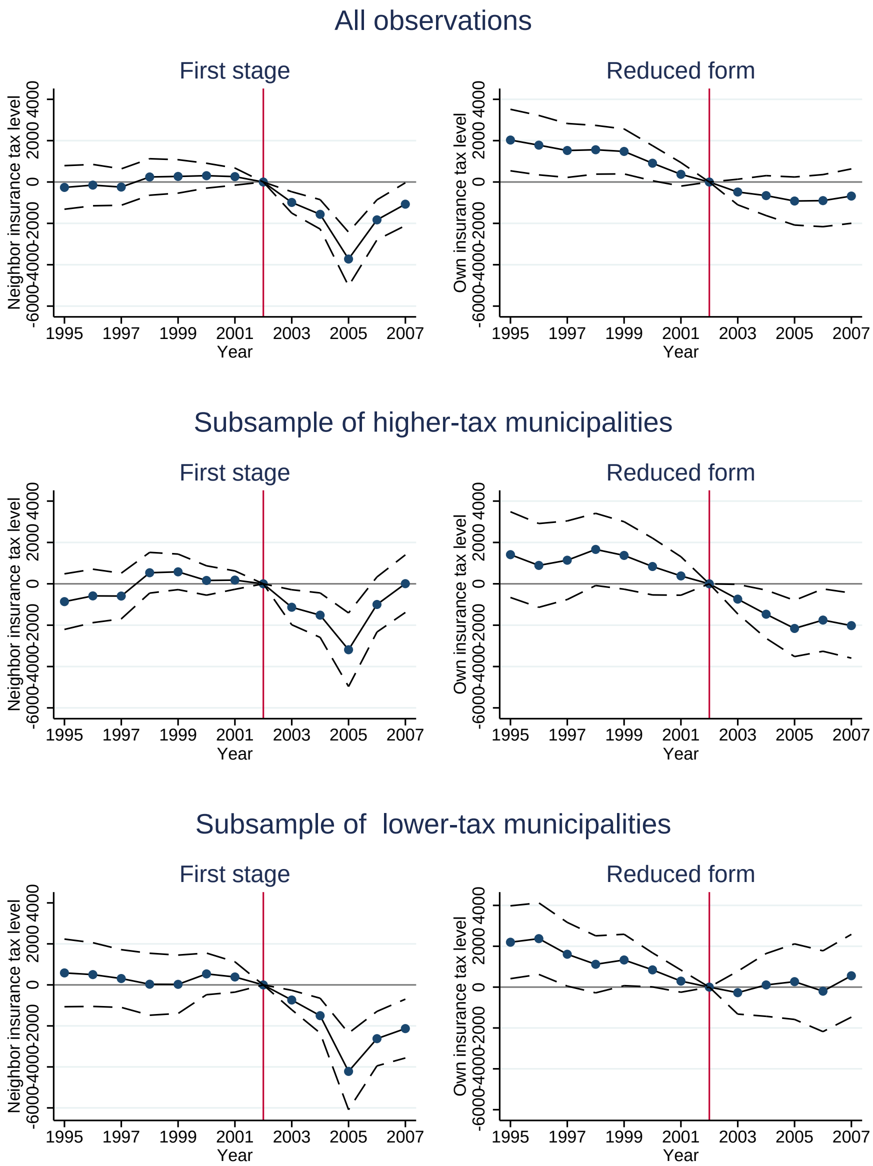
Tiebout, Charles M (1956) "A Pure Theory of Local Expenditures," *The Journal of Political Economy*, pp. 416–424.

Weingast, Barry, Kenneth A. Shepsle Shepsle, and Christopher Johnsen (1981) "The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics," *Journal of Political Economy*, Vol. 89, pp. 642–664.

Appendices

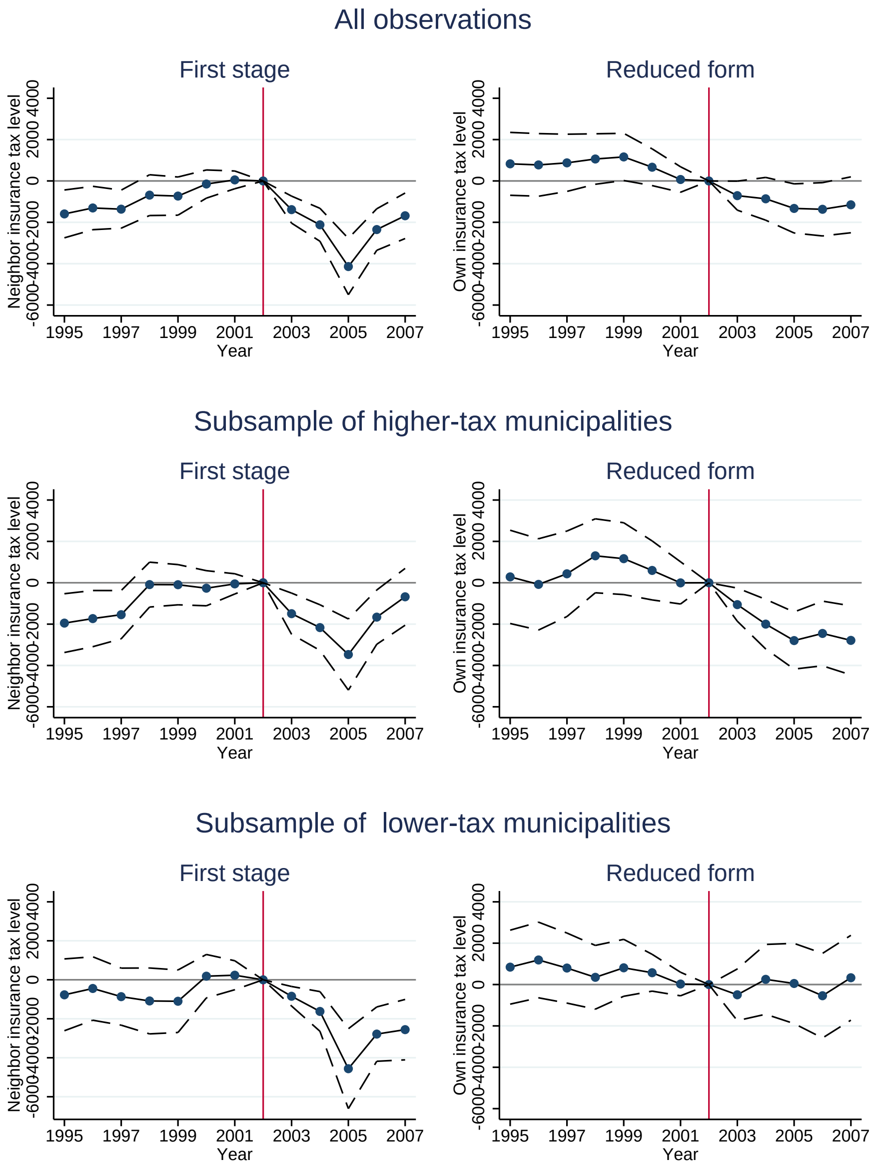
A Additional placebo results

Figure A.1: Estimates based on TSLS models without covariates and individual trends



Note: These estimation results are based on TSLS models (8) and (9), but the terms of own covariates, neighbor covariates and individual linear trends have been eliminated. Solid lines with dots present point estimates and dashed lines provide 95% confidence intervals. The confidence intervals are calculated with robust standard errors clustered by municipality.

Figure A.2: Estimates based on TSLS models with covariates but without individual trends



Note: These estimation results are based on TSLS models (8) and (9) but the terms of individual linear trends have been eliminated. Solid lines with dots present point estimates and dashed lines provide 95% confidence intervals. The confidence intervals are calculated with robust standard errors clustered by municipality.