CIRJE-F-1064

Do Central Grants Affect Welfare Caseloads? Evidence from Public Assistance in Japan

Masayoshi Hayashi
The University of Tokyo

September 2017; Revised in September 2018

Do Central Grants Affect Welfare Caseloads? Evidence from Public Assistance in Japan*

Masayoshi Hayashi**

Abstract: This study examines the existence of a "loosening effect" on a local social policy of central funding by exploiting a specific mechanism in the Japanese system of central grants and two institutional changes in categorical grants for local social assistance programs. Despite performing various types of estimations, we fail to substantiate the existence of the loosening effect. This result indicates the dominance of central control over local discretion in public assistance programs, and suggests that important factors are administrative rather than fiscal.

Keywords: welfare caseloads, social assistance, fiscal transfers, Japan
JEL codes: H73, H75, H77

---

* The final version of this discussion paper is published as "Do Central-Government Grants Affect Welfare Caseloads? Evidence from Public Assistance in Japan." *FinanzArchiv: Public Finance Analysis* 75(2), 152–186, 2019, DOI: 10.1628/fa-2019-0001.

** The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan (hayashim@e.u-tokyo.ac.jp).
I would like to thank Ronnie Schöb and an anonymous referee for their helpful observations and constructive suggestions through the reviewing process. This study is an outcome of a project begun more than 10 years ago. During the course, I have benefited from comments by my colleagues at different institutions, participants at various seminars and conferences, and readers of different versions of this manuscript. Among them, I am especially grateful to Michihito Ando, Shun-ichiro Bessho, Daiji Kawaguchi, Chul-In Lee, Akihiko Matsui, Haruko Noguchi, Fumio Ohtake, Hikaru Ogawa, and Heinrich W. Ursprung. The usual caveat applies. I also gratefully acknowledge the financial support from the JSPS Grants-in-Aid for Scientific Research (KAKENHI), Grant Numbers 20330064 and 15H01950.

1. Introduction

In countries where local (or subnational) governments implement welfare programs, the central government usually provides them with funds through a system of intergovernmental transfers. However, central and local governments often disagree over the design of such systems and have conflicting views on the effects of central funding. The central government may typically argue that more central grants would make local governments lenient in assessing the needs of welfare applicants, leading to an excessive number of welfare recipients and thus caseloads. By contrast, local governments, if they implement the programs set by the center, would contend that they are simply following the rules and that changes in the grants would not change their behavior.

However, while many studies have empirically examined the various determinants of welfare caseloads (e.g., Ziliak et al., 2000; Blank, 2001),[1] few have explored the effects of central funding. Indeed, while researchers have investigated the effects of central grants on local welfare *spending* (Chernick, 1998; Ribar and Wilhelm, 1999), their findings do not reveal the effect on *caseloads* if localities can set benefit levels as the states in the United States or provinces in Canada do. To deal with this issue, Baicker (2005) disentangle the effects on benefit level and caseload size. However, it is more straightforward to examine the case of a country where local governments are prohibited from changing the benefit levels.

In addition, identifying the effects of central funding is challenging for two main reasons. First, the structure of central grants provides an obstacle. For example, some grants have identical matching rates. Although changes may occur in the matching rates, they contain noise if a shock occurs nationwide when the rate alters. Second, while there may be variations

---

[1] Such research was started by non-US studies in the late 1970s (Spindler and Gilbreath, 1979; Gustafsson, 1984) and was stimulated by increased interest following the US welfare reform in the 1990s (Schiller and Brasher, 1993; Johnson et al., 1994; Schiller, 1999; Ziliak et al., 2000; Blank, 2001; Huang et al., 2004). Recent studies include Ayala and Pérez (2005), Page et al. (2005), Andini (2006), Cadena et al. (2006), Suzuki and Zhou (2007), Danielson and Klerman (2008), Hill and Murray (2008), Kneebone and While (2009), Jagannathan (2011), Snarr (2011), Berg and Gabel (2015), Klerman and Danielson (2016), and Ayala and Cano (2017).

in grant disbursements, endogeneity occurs if we use their amount as a regressor. For example, if the grants are matching, there exists reverse causation from local choices to the grant amounts. In addition, unobservable factors may correlate with both the grant amounts and the choices of local government (Holtz-Eakin, 1986). To tackle these issues, a growing number of studies rely on country-specific institutional mechanisms. For example, Baker et al. (1998) use the change from open- to closed-ended matching of federal grants in Canada, while Gordon (2004) uses the decennial changes in the parameters of a federal scheme that allocates grants for education in the United States. Others exploit the discontinuities in the equalization systems in Germany (Buettner, 2006) and Sweden (Dahlberg et al., 2008).

To examine the effect on welfare caseloads along this line of research, we exploit two features of the Japanese system of social assistance. First, local governments in Japan implement social assistance programs, called Public Assistance (PA), with the benefit levels firmly set by the center. As such, we can directly examine the effects on PA caseloads, since localities can only adjust their PA caseloads. While caseworkers supposedly follow the nationally uniform standards to assess the eligibility of PA applicants, anecdotal evidence indicates that they have some degree of discretion when applying such rules. For example, they may develop informal procedures to limit assistance to the entitled when their workload is high (Kobayashi, 2014). By contrast, when more resources are available, they might help applicants at the margin of eligibility.[2]

Second, the interplay of central transfers allows us to identify the effects of central funding. The central government disburses two types of grants to local governments for the implementation of PA programs. One is the Central Government Subsidy for Public Assistance (CGS-PA), which pays out a fixed proportion of PA benefits. All localities thus face an identical matching rate for their PA spending with no cross-section variations in each fiscal year (FY). The other is the Local Allocation Tax (LAT) that, albeit labeled a "tax," is a

[2] Because no active labor market programs were available in the years we investigate, locality could not adjust the caseloads through such programs.

general-purpose grant from the center. Only localities with "weak" fiscal capacity receive LAT grants. As we see later in detail, changes in LAT grants offset changes in CGS-PA disbursements. Obviously, such an offset does not happen to localities that do not receive LAT grants, allowing us to identify the changes in central funding for PA programs with the LAT recipiency status of localities. Moreover, since the inception of the current PA programs, the central government has changed the CGS-PA matching rate twice, with a reduction in FY1985 and an increase in FY1989. Given these institutional changes, we can exploit the interplay between the LAT and CGS-PA to identify the effect of central funding on local PA programs.

In addition to these institutional features, Japan is representative of the central–local policy debate mentioned at the outset. The center has traditionally displayed a strong aversion to fully funding local welfare programs, claiming that doing so would make localities spend welfare benefits excessively (Okuno, 1944). Indeed, this claim was so influential that the current CGS-PA programs only allow for partial central funding (Kasai, 1978). Furthermore, the center has been attempting to offload its costs of PA programs onto localities by reducing the matching rate, claiming that more central funding would "loosen" local welfare payouts, a claim constantly refuted by local governments. While the center successfully reduced the rate in FY1985, it failed to repeat the feat in FY2005 (Kimura, 2006), and this issue still occupies an important place in policy dialogs on intergovernmental fiscal relations in the country. We thus evaluate the unsubstantiated claim of the presence of such a "loosening effect" made frequently throughout the history of Japanese social policy.

We organize the rest of this paper as follows. Section 2 briefly introduces the Japanese PA system and intergovernmental transfers and elaborates on the identification strategies. Section 3 then sets out the estimation procedures and describes the data used for the estimation. Section 4 provides a descriptive analysis. Section 5 presents the results and discusses their implications, and section 6 concludes.

2. Institutional Mechanism and Identification

2.1 Institutional Background

Unlike the TANF programs in the United States, PA programs cover all types of households, including the elderly, single mothers, disadvantaged, and injured/sick who are judged unable to earn an income above the so-called "standardized" cost of living (SCL). Since the central government produces nationally standardized schedules to calculate the SCL, this allows for regional differences in consumer prices as well as in the characteristics of recipient households, including the number, age, sex, and health conditions of household members. The PA benefits are then set as the amount of the SCL that exceeds what an individual can earn with his/her best effort.

The system of local public administration in Japan consists of two tiers, with municipalities (cities, towns, and villages) as the first tier and prefectures as the second. As required by Japanese law, the welfare offices established in cities and prefectures implement social programs including PA.[3] Caseworkers at these welfare offices conduct a means test to assess the eligibility of PA applicants, following the procedures set out by the central government. Those eligible are required to fully exhaust their available resources, including financial support from family and relatives, as defined by the Japanese Civil Code. The benefits are then provided only if such income and resources are insufficient to cover the SCL. As such, benefits are supposed to be provided only after a careful examination of the financial situation of applicants.

While caseworkers are supposed to follow national standards, local discretion does arise. For example, they may limit assistance to the entitled by developing informal procedures within their individual welfare offices (Kobayashi, 2014). By contrast, if they obtain additional budgets from the central government, they may want to help applicants that would not be

[3] While towns and villages are not required to do so, a small number have chosen to set up their own welfare offices to implement social programs including PA. Prefectural welfare offices are responsible for providing PA to those residents in towns and villages that do not have their own welfare offices.

supported if the national standards were strictly applied. It was indeed this type of discretion to which the central government referred when it tried to reduce central funding for PA programs (Kimura, 2006). Note that since local governments did not have active labor market programs as a policy option in the years we consider, they could not adjust their caseload sizes through the welfare-to-work programs typical in most western countries.

As noted earlier, the CGS-PA and LAT are the two sources of central funding for PA programs. The CGS-PA comes from the budget of the Ministry of Health, Labour and Welfare, currently covering 75% of PA benefits. Meanwhile, the LAT is a general-purpose grant financed by national taxes along with other central revenue sources.[4] The LAT disbursement that a locality receives is the nonnegative difference between its Standard Fiscal Demand (SFD) and its Standard Fiscal Revenue (SFR), given as max{SFD – SFR, 0}. While the SFR estimates the local fiscal capacity of a given locality, the SFD estimates the level of expenditure required to maintain a "standard" quality of public services within that locality.

The crucial point in our analysis is that the SFD is the sum of standardized spending estimates for various expenditure categories. One such estimate is for PA spending, which we call "SFD-PA," consisting of (i) an estimate for PA benefits ($\bar{B}$) not covered by an estimated amount of the CGS-PA ($s\bar{B}$) where s is the CGS-PA matching rate, and (ii) an estimate for other PA expanses ($\bar{O}$) that include the running costs of welfare offices but exclude PA benefits. Note that the amount of the SFD-PA, $(\bar{B} - s\bar{B}) + \bar{O}$, does not match the annual PA expenses, as the values of $\bar{B}$ and $\bar{O}$ are predetermined for a given fiscal year. Meanwhile, note also that a change in the matching rate Δs alters the amount of the SFD-PA, and therefore, the SFD, by $-\bar{B}\Delta s$. This in turn causes a change in LAT grants by the same amount ($-\bar{B}\Delta s$) in localities that receive LAT grants (LAT localities). In contrast, for obvious reasons, this change does not affect the budgets of localities that do not receive LAT grants (non-LAT localities).

[4] The LAT consists of the Ordinary Local Allocation Tax and the Special Local Allocation Tax. In this study, the LAT refers to the former, which accounts for 96% of all LAT disbursements.

In FY1985, when the central government reduced the CGS-PA matching rate from 80% to 70%, it only compensated LAT localities for this reduction by increasing their LAT grants by $0.1\bar{B}$. Similarly, when the center increased the matching rate to its current value of 75% in FY1989, it only offset this increase for LAT localities by reducing their LAT grants by $0.05\bar{B}$. In other words, non-LAT localities suffered more from the 1985 rate reduction and benefitted more from the 1989 rate increase than LAT localities did.[5] Note that the effect of the difference is relative, that is, non-LAT localities are less (more) supported by the central grants than LAT localities are in the 1985 (1989) change. If a loosening effect exists, non-LAT localities would tend to attain smaller (larger) PA caseloads than LAT localities would in the 1985 (1989) change.

2.2. Effects of the CGS-PA Matching Rate Change

The argument above suggests that only non-LAT localities faced an increase (or a decrease) in own burden for PA programs. However, the mechanism is more complicated than it appears, as a change in the CGS-PA matching rate affects the relative price between PA benefits and other PA expenses for all localities, while the offsetting change in LAT grants only affects the general revenues of LAT localities. To obtain an appropriate perspective on the effect of the non-recipiency of LAT grants, we consider a simple model of local government choice which concerns the number of PA recipients or PA caseloads (Y), other PA expenses (O), and local government expenses other than PA (Z). As the central government sets the benefit level, we regard per recipient benefits (b) as exogeneous and given to localities. Without loss of generality, we set per recipient benefits as unity $b = 1$, so that total PA benefits (B) equal PA caseloads, $B = bY = Y$. Local revenue thus consists of CGS-PA payments ($sB = sY$), local taxes (T), and LAT grants (G) for LAT localities, but only CGS-PA payments (sY) and local taxes (T) for non-LAT localities. Therefore, the budget constraints are given as

[5] As the SFD-PA is an estimate, the central government could not exactly compensate LAT-receiving localities. However, the estimates performed very well, as Nakai (1988) confirms in examining the offsetting changes in the SFDs, for cities in Osaka prefecture in the 1985 reduction.

$$(1-s)Y+O+Z=T+G \tag{1}$$

for LAT localities and

$$(1-s)Y+O+Z=T \tag{2}$$

for non-LAT localities.

When the matching rate (s) is fixed, LAT grant (G) in Eq. (1) is independent of the actual volume of PA caseloads (Y), as the SFD-PA is based on predetermined estimates. Meanwhile, when the matching rate changes by Δs, the price of Y, i.e., $1-s$, changes for both LAT and non-LAT localities by $-\Delta s$. In addition, this change also alters G such that

$$\Delta G=-\bar{Y}\Delta s \tag{3}$$

where $\bar{Y}$ is an estimate of Y. Therefore, in addition to the price change, a LAT locality with Eq. (1) faces a lump-sum increase in revenue when the matching rate decreases ($\Delta s < 0$) or a lump-sum decrease in revenue when the rate increases ($\Delta s > 0$).

Assume that a locality has an objective function $V(Y, O, Z)$ with the standard property, and freely chooses the three "goods." We see from the arguments above that a decrease (increase) in s raises (reduces) the "price" of PA caseloads both in LAT and in non-LAT localities, but increases (reduces) only the budget of LAT localities. Therefore, if PA caseloads are a normal good, LAT and non-LAT localities decrease (increase) their caseloads in response to a reduction (rise) in the matching rate. However, the caseload reduction (increase) is larger (smaller) in non-LAT localities than in LAT localities. Therefore, if the matching rate decreases (increases), the difference between the PA caseload change in LAT and in non-LAT localities, is negative (positive) when using LAT localities for comparison.

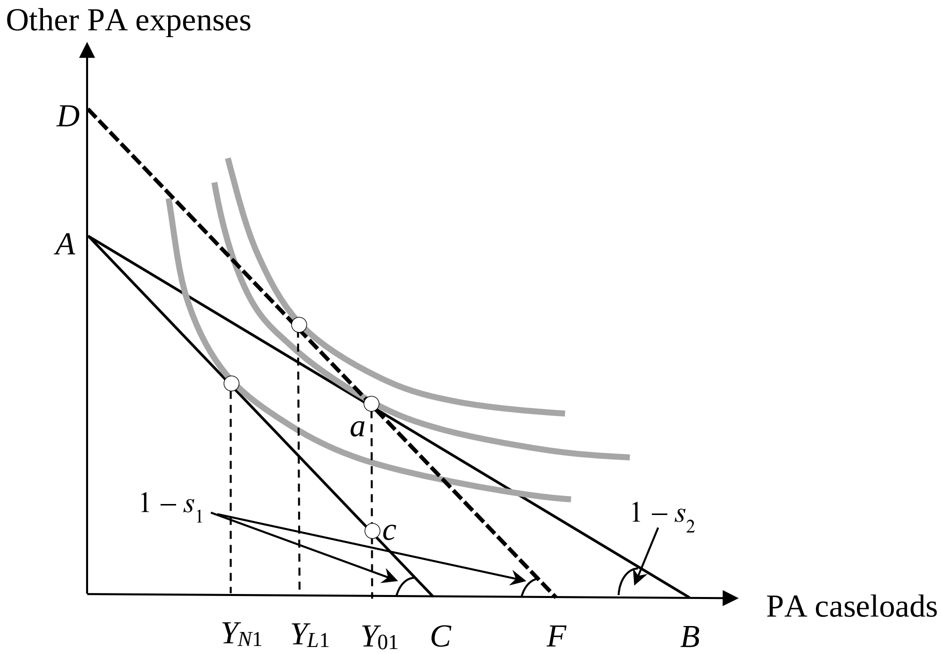
The two panels in Figure 1 describe such an effect. Assuming the weak separability of $V(\cdot)$ between (Y, O) and Z, they illustrate the choice between PA caseloads (Y) and other PA expenses (O) through two-stage budgeting (e.g., Deaton and Muellbauer, 1980). We assume that all goods are normal, and that localities are identical except for their LAT status. Panel A describes the effects of the matching rate reduction ($s_1 \rightarrow s_2$, $s_1 > s_2$). With the initial budget line AB, both LAT and non-LAT localities choose their caseloads at the same level, Y_{01}. The rate reduction rotates the budget line from AB to AC, on which a non-LAT locality chooses its caseloads at Y_{N1}. If $\bar{Y} = Y_{01}$, the change also increases the LAT grant by *ac*. This makes a LAT locality face a different budget line DF, which parallels AC, and choose its caseloads at Y_{L1}. Thus, the difference in differences (DD) for this rate reduction is negative when the LAT locality serves as the reference as $(Y_{N1} - Y_{01}) - (Y_{L1} - Y_{01}) = Y_{N1} - Y_{L1} < 0$.

Meanwhile, Panel B illustrates the analogous but opposite case of a rate increase ($s_2 \rightarrow s_1$, $s_1 > s_2$), starting now with the budget line AC where both localities choose Y_{02} of their caseloads. An increase in the rate rotates the budget line from AC to AB, on which a non-LAT locality chooses its caseloads at Y_{N2}. Since the change now reduces the LAT grant by *de* if $\bar{Y} = Y_{02}$, a LAT locality faces a different budget line HJ, parallel to AB, where it selects its caseloads at Y_{L2}. Here, the DD for this rate increase is positive when the LAT locality is the reference since $(Y_{N2} - Y_{02}) - (Y_{L2} - Y_{02}) = Y_{N2} - Y_{L2} > 0$.

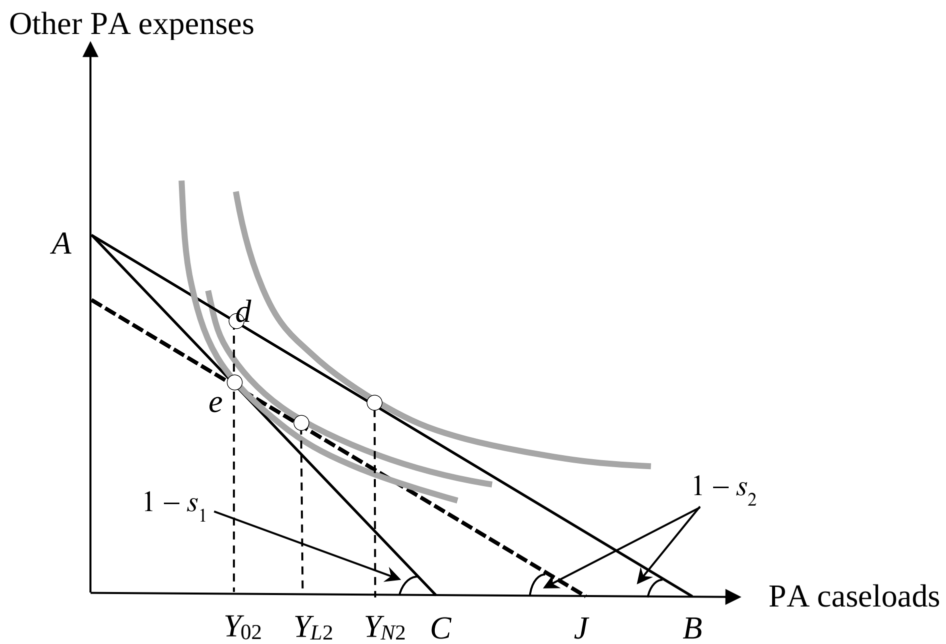
We could cast doubt on the assumption that explains the two panels in Figure 1, arguing that it does not reflect the actual process through which PA eligibility is assessed. Formally, caseworkers are supposed to follow nationally uniform procedures that are independent of local fiscals. A hierarchy of supervision and audit procedures is in place to ensure that PA programs are implemented according to the rules. Caseworkers are supervised by directors at welfare offices, who are then supervised by managers in the relevant welfare sections of local governments. In addition, if municipalities implement the program, they are audited by prefectures. If localities follow the national rules, they have to provide PA benefits to anyone who satisfies the uniform eligibility criteria. In other words, Y in Figure 1 is not the variable they can choose but rather the parameter they have to take as given. In other words, choices are only made over O and Z, accepting Y as independent of their budget and preferences.

The claim of the existence of a loosening effect denies the expected effects of these audits and supervisions since—if they are effective—PA caseloads should not be affected by local fiscal factors including the size of central funding. The results of the following estimation therefore constitute indirect evidence that shows whether the PA system is implemented according to what the system ostensibly expects.

Figure 1 *The effects of the matching rate changes*



Panel A: The case of a rate reduction



Panel B: The case of a rate increase

3. Models and Estimation Methods

We take advantage of these two changes in the CGS-PA matching rate along with the LAT status in those years to identify the effect of central grants on PA caseloads. We use the difference-in-differences (DD) design to examine the effects of central government grants on PA caseloads. The outcome variable is a measure of PA caseload size at the municipal level. Following Huang et al. (2004) and Danielson and Klerman (2008), we measure caseloads in natural logarithms.[6] For the estimation, we divide our sample of local governments into two groups (LAT and non-LAT localities) and obtain sets of DD estimates. Non-LAT localities serve as the treatment group. The institutional changes we exploit are the matching rate decrease in FY1985 (from 80% to 70%) and increase in FY1989 (70% to 75%). In what follows, we use two sets of estimations. The first set is "canonical" in the sense that it adopts a pair of periods before and after the change. Meanwhile, the other set adopts panel data that contain annual data on every year from 1981 to 1992.

3.1 Estimation with two-period data

For the DD estimation that adopts a pair of periods before and after the change, we use the year immediately before the year of the change as the base period t_B (t_B = 1984 for the 1985 change; t_B = 1988 for the 1989 change) and one of the following three alternatives as the end period: t_A, $t_A = t_B + 1$, $t_A = t_B + 2$, or $t_A = t_B + 3$ (t_A = 1986, 1987, or 1988 for the 1985 change; t_A = 1990, 1991, or 1992 for the 1989 change). By employing these pairs of periods, we use the following four DD estimators. The first is the baseline DD estimate:

$$\hat{\delta} = \frac{1}{n_k}\sum_{j}^{n_k} \left(\ln Y_{kt_A} - \ln Y_{kt_B}\right) - \frac{1}{n_l}\sum_{l}^{n_l}\left(\ln Y_{lt_A} - \ln Y_{lt_B}\right), \tag{4}$$

where Y_{it} is PA caseload size in locality i in year t, subscripts k and l indicate units in the treatment and control groups, and n_k and n_l are the number of units in each group.

---

[6] Another index is the logarithm of the ratio of caseload size to population (e.g., Blank, 2001). However, our measure is more useful in this context as its difference is interpretable as a growth rate. In addition, the DD regression can allow for the ratio index by including the log of population as a regressor.

We can obtain this estimate as an OLS estimate for δ from the regression model:

$$\Delta\ln Y_i = \gamma + \delta \cdot D_i + e_i, \tag{M1}$$

where $\Delta\ln Y_i \equiv \Delta\ln Y_{it_A} - \Delta\ln Y_{it_B}$, γ is a drift, D_i is the treatment (dummy) variable for non-LAT localities, and e_i is an error term. Note that (M1) can be regarded as the differenced form of the following regression model:

$$\ln Y_{it} = \alpha_i + \gamma_t + \delta \cdot D_{it} + u_{it} \tag{5}$$

for $t = t_B$ and t_A, where α_i is unobserved heterogeneity, $\gamma_{t_A} = \gamma_{t_B} + \gamma$, D_{it} is the treatment (dummy) for non-LAT localities in year t_A, such that $D_{it_A} = D_i + D_{it_B}$ with $D_{it_B} = 0$ for all observations by definition, and u_{it} is an error term such that $u_{it_A} = u_{it_B} + e_i$.[7]

The DD estimate is valid only if the parallel trend assumption holds. As parallel trends are not necessarily warranted in (M1), we could control for different trends, if any, by including additional regressors. A typical way to do this is to add linear or quadratic time trends, which is not feasible in the current case, as we only use two-period data before and after the change ($t = t_B$ and t_A). Instead, we may add a vector of covariates $\boldsymbol{X}_i$ into (M1), anticipating that this inclusion controls for different trends among cities. Thus, as the second set of our DD estimates, we use the OLS estimate of δ in the following regression model:

$$\Delta \ln Y_i = \gamma + \delta \cdot D_i + \boldsymbol{X}_i'\boldsymbol{\theta} + e_i. \tag{M2}$$

Analogous to the relation between (M1) and Eq. (5), notice that (M2) is considered to be the differenced form of the following model for $t = t_B$ and t_A:

$$\ln Y_{it} = \alpha_i + \gamma_t + \delta \cdot D_{it} + \boldsymbol{X}_i'\boldsymbol{\theta}_t + u_{it}. \tag{6}$$

Note that $\boldsymbol{X}_i$ is observed in year t_A or some earlier period, and the vector of its coefficients $\boldsymbol{\theta}_t$ is assumed to take different values in t_A and t_B such that $\boldsymbol{\theta}_{tA} = \boldsymbol{\theta}_{tB} + \boldsymbol{\theta}$. The latter assumption captures the different effects of $\boldsymbol{X}_i$, reflecting the differences in time from its observation.

---

[7] With the two-period data ($t = t_B$ and t_A), the OLS estimate of δ in (M1) or (M2) is identical to the within estimate of δ in Eq. (5) or Eq. (6), respectively. See Wooldridge (2010, pp. 321–335) for more details.

Instead of parametrically including the covariates and unobserved heterogeneity as in (M2), we can allow for such heterogeneity nonparametrically by using the kernel propensity score matching DD (PSM-DD) estimator of Heckman et al. (1997). This estimator is obtained for our case as

$$\hat{\delta}_{PSM} = \frac{1}{n_k}\sum_{k}^{n_k}\left(\ln Y_{kt_A} - \ln Y_{kt_B}\right) - \frac{1}{n_k}\sum_{k}^{n_k}\sum_{l}^{n_l} w_{kl}\cdot\left(\ln Y_{lt_A} - \ln Y_{lt_B}\right), \tag{M3}$$

where w_{kl} is a weight for a pair of k in the treatment group and l in the control group, and the other legends are the same for (M1). Note that weight w_{kl} is given as

$$w_{kl} = \frac{K\left(\frac{p_k - p_l}{h}\right)}{\sum_l K\left(\frac{p_k - p_l}{h}\right)} \tag{7}$$

where $p_i = p(\boldsymbol{X}_i)$ is the estimated propensity score, $K(\cdot)$ is the kernel function, and h is the value of the bandwidth. We can also obtain another set of PSM-DD estimates by restricting our sample of observations to those that yield propensity scores that fall in the overlapping ranges of the propensity scores for the treated and untreated units. This estimator with common support arguably increases its internal validity. We use this PSM-DD with common support as the fourth set of our DD estimates (M4).

Inferences with the DD estimation above are complicated by the severe underestimation of the standard errors if the residuals exhibit some form of correlation. One of the most common practices to handle this issue is to employ the cluster-robust variance estimator at the group level. Since local policies within a given prefecture are likely to be highly correlated, we cluster the standard errors at the prefecture level for (M1)–(M4).

3.2 Estimation with more than two-period data

When estimating (M2), we intend to control for the different trends among the groups of cities by including as regressors covariates $\boldsymbol{X}_i$ in (or before) the year of change. Alternatively, we could augment the regression model with the polynomial trends for a specific group of observation units, along with year fixed effects. While this is a standard method in the literature (Friedberg, 1999; Autor, 2003; Besley and Burgess, 2004; Angrist and Pischke, 2009), we could not use it with (M2), as we need at least three periods of data to identify the time trends. In addition, since the 1985 and 1989 changes are only separated by three years, it would be of interest to estimate the effects of the two changes jointly rather than separately, as in the estimations of (M1)–(M4). Therefore, we estimate variations of the following generic model, using panel data spanning every year from FY1981 to FY1992:

$$\ln Y_{it} = \sum_{\tau} \delta_{\tau} \cdot D_{i,\tau} + \alpha_i + \gamma(i.t) + g_i \cdot \sum_{p=1}^{P} \phi_p \cdot t^p + \sum_{q=1}^{Q} \rho_q \cdot \ln Y_{it-q} + \epsilon_{it} \qquad \text{(M5)}$$

where t = 1981, 1982, …, 1992, τ = 1985 or 1989, δ_τ is the effects of the change in year τ, $D_{i,\tau}$ is the treatment variable, which takes unity for non-LAT localities in year τ and after (and zero otherwise), and ϵ_{it} is the error term. We elaborate on the other elements as follows.

First, $\gamma(i, t)$ is the fixed effect that takes a specific value in year t, of which we consider two versions. First, we assume that $\gamma(i, t)$ is a typical fixed time effect whose values are common to all localities but different over the years, $\gamma(i, t) = \gamma_t$. Thus, a sequence of $\{\gamma_t\}$ forms annual trends common for all cities in the sample, including both LAT and non-LAT localities. Second, we relax this common-effect assumption to allow $\gamma(i, t)$ to take a value from among prefectures, $\gamma(i, t) = \gamma_{jt}$ with j indexing the prefecture for locality i. We can estimate a set of γ_{jt} by including the interactions between year and prefecture dummies in the regression model. Obviously, these prefecture-year effects control for factors whose effects are identical within a prefecture but different among prefectures. We may substitute the prefecture-year effects for covariates $\boldsymbol{X}_i$, especially if the covariates are highly correlated within a prefecture.[8]

Second, $g_i \cdot \sum_{p=1}^{P} \phi_p \cdot t^p$ is a time trend for the treated units (non-LAT localities) that deviates from the fixed year effects (γ_t or γ_{jt}), where g_i is a dummy variable for non-LAT

[8] This is a convenient property, as the data for $\boldsymbol{X}$ used in (M2), being obtained from the national census (surveyed every five years), are not available for every year.

localities and ϕ_ps are parameters to be estimated. If the trends differ between treated and untreated units, we expect these parameters to capture the differences. We may also regard this inclusion as an additional substitute for allowing for the effects of $\boldsymbol{X}$. We consider two cases, one with a linear time trend (P = 1) and the other with a quadratic time trend (P = 2).

Lastly, $\sum_{q=1}^{Q} \rho_q \cdot \ln Y_{it-q}$ are dynamic effects where ρ_qs are autoregressive parameters. We consider four cases: one with no dynamic effects ($\rho_q = 0 \ \forall q$) and three others with Q = 1, 2, and 3. The dynamic effects, or the effects from the lagged dependent variables, may be relevant if adjustments of PA caseloads are costly, and thus, "partial" (Nerlove, 1956). This would also be the case if the workload of caseworkers affects the current volume of caseloads.[9]

These patterns of the three elements yield 16 specifications of (M5). For the four specifications that have no dynamic effects, we can validly estimate their within-transformed versions (i.e., the deviations from the time means for a given unit) by OLS. However, such OLS estimation does not yield consistent estimates when the model includes lagged dependent variables as regressors.[10] Therefore, to estimate the specifications of (M5) that have dynamic effects, we adopt the standard GMM (generalized method of moments) estimators for dynamic panel estimation. Such estimators typically consist of four types: the one- and two-step difference estimator (Holtz-Eakin et al., 1988; Arellano and Bond, 1991) and the one- and two-step system estimator (Arellano and Bover, 1995; Blundell and Bond, 1998). We adopt all four to estimate the 12 variations of (M5) with $Q \geq 1$,[11] using the instruments that include the relevant lags of $\Delta \ln Y_{it-q}$.[12]

---

[9] This observation was made by the reviewer of this paper.

[10] This is because the within-transformed lagged regressor ($\ln Y_{it-q} - T^{-1}\sum_s \ln Y_{is-q}$) is correlated with the within-transformed error term ($\varepsilon_{it} - T^{-1}\sum_s \varepsilon_{is}$) when the number of time-series observations T is small. Similarly, the OLS estimation of the differenced versions of (M5) does not yield consistent estimates either, as the differenced lagged regressor $\Delta \ln Y_{it-q} \equiv \ln Y_{it-q} - \ln Y_{it-q-1}$ is correlated with the differenced error term $\Delta\varepsilon_{it} \equiv \varepsilon_{it} - \varepsilon_{it-1}$. See Nickell (1981) and Wooldridge (2010, pp. 371–374)

[11] We obtain all the dynamic panel estimates using a STATA module (xtabond2) written by Roodman (2009a). When using the two-step system estimator, we adjust the variance-covariance matrices with the finite sample correction presented by Windmeijer (2005).

[12] The choice of instruments is standard, as proposed by Holtz-Eakin et al. (1988). Such instruments may not be valid if serial correlation exists in the error term of the dynamic panel models, about which Angrist and Pischke (2009) raise serious concerns. One reason to estimate the models with Q = 2 or 3 is to allow

4. Implementation and Descriptive Analysis

4.1 Samples and PA Caseload Data

We use data on 254 cities in 11 prefectures (of the 47 prefectures of Japan) in the 1980s and early 1990s (1981−1992). We select cities in prefectures that we consider to be more "urban" than others, which consist of Ibaraki, Tochigi, Saitama, Chiba, Tokyo, Kanagawa, Shizuoka, Aichi, Kyoto, Osaka, and Hyogo.[13] We make this choice since LAT localities are concentrated in the other areas and the inclusion of cities in those areas into the sample only leads to a disproportionate increase in the size of the control group. Furthermore, it may also increase the heterogeneity in the sample since the socioeconomic characteristics of cities in "urban" and "non-urban" areas may differ. Note also that the sample excludes the two cities (Chiba and Sakura) whose administrative boundaries changed during the study period. We do not consider this to be a serious selection problem since cities rarely change their boundaries for reasons related to their PA programs.

PA caseloads are the number of households that received PA benefits. The main sources for PA caseload data at the municipal level for the 1980s and early 1990s are prefectural statistical yearbooks. Since each prefecture independently compiles such data to publish in its statistical yearbook, the formats of their data are not necessarily uniform. Several prefectures (e.g., Shizuoka prefecture) do not list PA data at the city level. We therefore obtain data for the cities in Shizuoka from the System of Social and Demographic Statistics (SSDS) compiled by the Statistical Bureau in Japan.[14]

---

for these concerns, as we could reasonably attenuate the serial correlation by augmenting the model with further lagged dependent variables ($\ln Y_{it-q}$, $q > 1$). Another issue with the standard choice of instruments is that it may result in "too many instruments" (Roodman, 2009b). However, our number of excluded instruments will be between 64 and 80 depending on the model specifications, which is well below the number of cross-section observations, 254, satisfying the rule of thumb that the maximal number of instruments should not exceed the number of cross-section observations.

[13] We intended to include the data for cities in Gunma prefecture, as they may also be regarded as "urban." However, their caseload data are unavailable for the years we examine herein.

[14] The data are not publicly accessible in the sense that we had to purchase them from Infonica, which commercially distributes the data set from the SSDS.

The 11 prefectures record municipal PA caseload data with different timings. Kanagawa and Osaka document their municipal caseload data on the last day of an FY, while Hyogo records them on the first day of an FY. We therefore take a one-year lag of the caseload data for the cities in Hyogo prefecture. The other eight list their data as yearly averages. While these differences are not ideal, the DD estimation should mitigate their adverse effects since it differences the data along the timeline, which allows for unobserved heterogeneities including the different timing of recordkeeping among prefectures.

4.2 Treatment and Control Groups

We use the non-receiving status of the LAT as the treatment variable for the institutional changes in FY1985 and FY1989. Note that the treatment effects on the two occasions point to different directions. In FY1985, the central government reduced the PA matching rate. Since non-LAT localities were not compensated through the LAT, the treatment (the non-recipiency of the LAT) indicates a reduction in the central funding. Meanwhile, the center increased the matching rate in FY1989. Since LAT localities faced corresponding reductions in their LAT grants in this year, only non-LAT localities enjoyed the benefits of this rate increase. That is, the treatment (non-recipiency) indicates an increase in the central funding.

However, it is difficult to uniquely identify the treated and untreated since there are a large number of "movers" (i.e., cities that flip their LAT status between LAT recipiency and non-LAT recipiency) from year to year. Because there is no single way of choosing a treatment that solves this issue, we employ several versions. For the estimation of (M1)–(M4), we categorize as the treatment group those cities whose non-receiving status did not change from t_B to t_A, which should be a natural choice to make. As mentioned in Section 3, we set $t_B = 1984$ and 1988 for the changes in FY1985 and FY1989, respectively, and examine the three patterns of $t_A = t_B + 1$, $t_A = t_B + 2$, and $t_A = t_B + 3$. The choice of the control group is not unique either. We consider the following two cases. For one, we include movers (i.e., cities with a changing LAT status) in the control group. For the other, we exclude movers from the control group.

The latter is "no-mover" case which restrict the sample only to those cities that maintained the either of the LAT status in every year from t_B to t_A.

There may be another issue for us to consider when we construct the treatment variable. Cities whose values of the SFR/SFD ratio are around unity may be uncertain of their LAT status in future years since their status could easily be reversed. This observation would yield a notion that the effect might reveal itself more conspicuously for cities having firmer expectations of their LAT status, which may result from a longer stream of constant LAT status. We may then put a stricter condition for selecting the treated units that imposes a longer constancy of non-LAT receipt. Although the choice of the length of this period is arbitrary, we pick it as one from $t_B - 3$ to t_A, and call it the "longer-horizon" case. Again, there is an issue with movers. We thus examine two longer-horizon cases: one that includes movers in the control group and the other that excludes them from the sample. Since the latter case excludes units whose values of the SFR/SFD ratio were volatile around unity, it compares two sets of cities *both* having firmer expectations of their LAT status. If any effect of central funding existed at all, it would therefore reveal itself even more noticeably in this restricted sample.

For the estimation of the 16 specifications of (M5), we analogously define the treatment units as those cities whose LAT recipiency status did not change in every year from 1981 to 1992, which places an even stricter condition than the longer-horizon case for (M1)–(M4). Analogously to the previous cases, we again examine two samples, movers in the control group and movers excluded. The latter compares two sets of cities *both* having even firmer expectations of their LAT status than the no-mover longer-horizon cases for (M1)–(M4).

4.3 Validity of the Treatment Variables

LAT status (or non-LAT status) might be subject to manipulation through the control by local governments of the SFD or SFR, or both, especially when their SFR/SFD ratios are close to unity. For example, if their ratios are just above unity when the center is about to reduce the PA matching rate, cities might try to reduce the ratio to be compensated by an increase in LAT disbursements. By contrast, if the ratios are just below unity when the matching rate increases, they might also try to increase their ratios since only non-LAT localities can enjoy the rate increases. Therefore, there might be incentives for cities whose SFR/SRD ratios are around unity to move to the other side of the threshold, depending if the CGS-PA increases or decreases.

If this is the case, self-selection is an issue. We therefore formally test the selection problem by exploiting a recent development in manipulation testing. Such tests generally examine whether there is a discontinuity in the density of observations at a known cutoff. Rejecting the null hypothesis of no discontinuity is interpreted as evidence of self-selection into the control and treatment groups. For this, we employ the manipulation test proposed by Cattaneo et al. (2017a), or the CJM test, based on a local polynomial density estimator that, they argue, improves both the size and the power relative to other tests.

To obtain a more precise density shape, we use three samples for the CJM test, which is larger than the sample used for our DD estimation below. They are two cross-section samples of all cities in FY1985 and FY1989 (N = 674) and a sample that pools all cities from FY1981 to FY1992 ($T \times N$ = 8,088). The variable over which the density is defined is the SFR/SFD ratio. The cutoff value is unity (1 = SFR/SFD) where cities alternate their LAT status. Table 1 lists the results of the CJM tests, showing the relevant test statistics for the three samples (FY1985, FY1989, and FY1981–1992) by using three kernel functions (uniform, triangular, and Epanechnikov), along with an effective number of observations used for each side of the cutoff value.[15] By providing relatively large p values (0.126–0.485), none of the test statistics reject the null hypothesis at standard levels of statistical significance. This then allows us to conclude that the manipulation of the LAT status is not an issue in our estimation.

[15] To obtain the test statistics, we use rddensity, a STATA module written by Cattaneo et al. (2018). Except for the choice of the kernel function, we use the default setting of the command for the parameters required to obtain the test statistics.

Table 1 *Manipulation tests*

	Kernel	Effective number of obs.		Test statistics	
		Left	Right	$T_3(h_2)$	p values
FY1985 ($N = 674$)	Uniform	167	77	−1.530	0.126
	Triangular	416	83	−0.745	0.456
	Epanechnikov	296	80	−1.352	0.177
FY1989 ($N = 674$)	Uniform	225	95	−0.925	0.355
	Triangular	257	95	−0.698	0.485
	Epanechnikov	247	95	−0.831	0.410
FY1981–92 ($TN = 8{,}088$)	Uniform	2,468	813	−0.901	0.368
	Triangular	2,014	756	−1.208	0.227
	Epanechnikov	2,266	764	−1.242	0.214

Notes: (i) ***: $p \leq 0.01$; **: $0.01 < p \leq 0.05$; *: $0.05 < p \leq 0.10$. No results are statistically significant above at these standard levels of significance, as the p values show. (ii) To obtain the test statistics, we used rddensity, a STATA command written by Cattaneo et al. (2017b). Except for the choice of the kernel function, we used the default setting of the command for the parameters required to obtain the test statistics, $T_3(h_2)$, where 3 is the order of the local polynomial for constructing the bias-corrected density point estimator and 2 is the order of the local polynomial for constructing the density point estimator. See Cattaneo et al. (2017b).

4.4 Covariates and Descriptive Analysis of PA Caseloads

Table 2 summarizes the descriptive statistics for the data on PA caseloads, along with those for the covariates used for (M2)–(M4). The covariates are measured at the city level and constructed from the data collected in the national census. They consist of (a) population (in natural logarithms), (b) the unemployment rate, (c) a poverty index,[16] and (d) average household size (in natural logarithms) as well as the shares of (e) single mother households, (f) single elderly households, and (g) the Korean population and share of the working population in (h) manufacturing and (i) services. Note that the national census is conducted every five years. In other words, we can only obtain annual values for 1985 and 1990 during the periods under our examination (i.e., between FY1981 and FY1992). As we may set their

[16] The poverty index is the share of local residents whose annual incomes are above the minimum taxable income (*Annual Report on Municipal Finances* compiled by the Ministry of Internal Affairs and Communication for FY1985) among the working population aged 16–65 (1985 national census).

values in the year when or before the change occurred, the only choice is to obtain data for these covariates from the 1985 national census.

Table 2 *Sample statistics*

	Mean	Std. Dev.	Minimum	Maximum	Unit
PA caseload FY1981	870.3	2981.2	46.0	35,621.0	000 cases
PA caseload FY1982	901.0	3,091.2	45.0	36,663.0	〃
PA caseload FY1983	934.4	3,199.7	38.0	37,545.0	〃
PA caseload FY1984	952.0	3,247.4	33.0	37,995.0	〃
PA caseload FY1985	950.4	3,210.0	32.0	37,051.0	〃
PA caseload FY1986	922.3	3,111.9	29.0	35,768.0	〃
PA caseload FY1987	890.7	3,009.0	25.0	34,366.0	〃
PA caseload FY1988	860.1	2,927.3	19.0	33,411.0	〃
PA caseload FY1989	826.6	2,832.6	19.0	32,541.0	〃
PA caseload FY1990	796.1	2,760.2	21.0	31,878.0	〃
PA caseload FY1991	778.8	2,730.2	21.0	31,698.0	〃
Households	56.7	112.7	6.5	1,027.0	000 households
Population	175.2	319.4	21.7	2,993.0	000 persons
Poverty index	0.254	0.043	0.140	0.398	n.a.
Unemployment rate	0.029	0.008	0.015	0.058	〃
Share of Korean residents	0.004	0.005	0.000	0.037	〃
Share of single mother households	0.013	0.003	0.008	0.024	〃
Share of elderly households	0.029	0.013	0.010	0.106	〃
Share of workers in the manufacturing sector	0.379	0.082	0.145	0.609	〃
Share of workers in the service sector	0.565	0.096	0.293	0.825	〃

Notes: (i) The sample size is 254. (ii) The values except for PA caseloads are taken from the 1985 national census. (iii) The sources of PA caseload data are mainly the statistical annals of the relevant prefectures.

Some of these variables may require elaboration. For the poverty index (c), we use the share among the working population (aged 16–65) of residents whose annual incomes are above the minimum taxable income. We then obtain the value of one minus this share as the poverty index, since those who do not pay income taxes usually earn no or little income. Single mothers (e), single elderly (f), and Korean residents (g) are well-accepted characteristics in

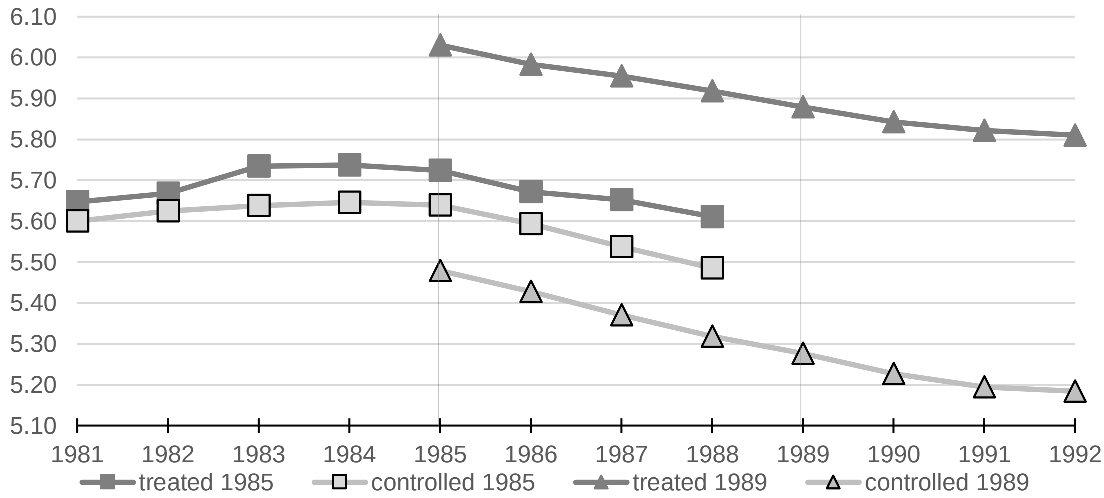
Japan that correlate with the receipts of PA benefits. The industrial shares of workers (h and i) should reflect regional characteristics, which may affect local employment opportunities.

The literature suggests that caseworkers may share a given collective value within their organizations, which exerts a major influence on their daily practices (Keiser and Soss, 1998). In addition, community values may discourage eligible individuals from applying and/or cause caseworkers to take tough positions on eligibility assessment (Grubb, 1984; Weissert, 1994). For example, in areas with strong family ties where residents may easily receive financial support from family and relatives, caseworkers tend to limit assistance to the entitled. In areas with traditional values, more people may feel stigmatized if they receive social assistance. In such areas, we may expect fewer PA applicants. Since these factors are likely to differ across cities but unlikely to change during a short period, we could conceptualize them as unobserved heterogeneity (α_i), which is allowed for in all cases of our estimation.

Before we perform the estimation, let us visually examine the trends of the average PA caseloads (in natural logarithms) of the treated and untreated before and after the treatments (i.e., changes in the CGS-PA). However, it is difficult, if not impossible, to graph all the averages of the treatment and control groups that correspond to the different definitions of the treatments, alternative methods of estimation, two institutional changes, and different sizes of the sample. Therefore, we instead show them only for two cases with longer horizons. As we argued above, the effect, if existed, might reveal itself more noticeably in such cases.

The first is the case with non-movers and a longer horizon for $t_A = t_B + 3$. Figure 2 depicts the average values of PA caseloads in natural logarithms for the treatment and control groups. Note that since the periods under consideration are different (FY1984–FY1997 for the 1985 change and FY1988−FY1991 for the 1989 change), the units in the treated and control groups also differ. With $n_k = 30$ and $n_l = 157$ for the 1985 change and $n_k = 67$ and $n_l = 132$ for the 1989 change, the values for the overlapping periods (FY1985, FY1986, FY1987, and FY1988) are different. For the 1985 change (FY1981−FY1988), the two averages changed almost parallelly from FY1983 to FY1986. After FY1987, the treated lifted slightly compared with the untreated, which implies that the treated had a less negative difference. This should then yield a positive effect, *which is opposite to that previously claimed*. On the contrary, for the periods for the 1989 change (FY1985−FY1992), the line for the treated became slightly flatter before FY1989, which is consistent with the post-FY1985 trends of the lines for the 1985 change. Meanwhile, the line for the untreated had a slightly larger dip after FY1989. This implies that the treated have a less negative difference than the untreated, yielding a positive effect. This is indeed consistent with the popular claim.

Figure 2 *Average PA caseloads (in logarithm) of the treated and controlled with different no-movers for the 1985 and 1989 changes*



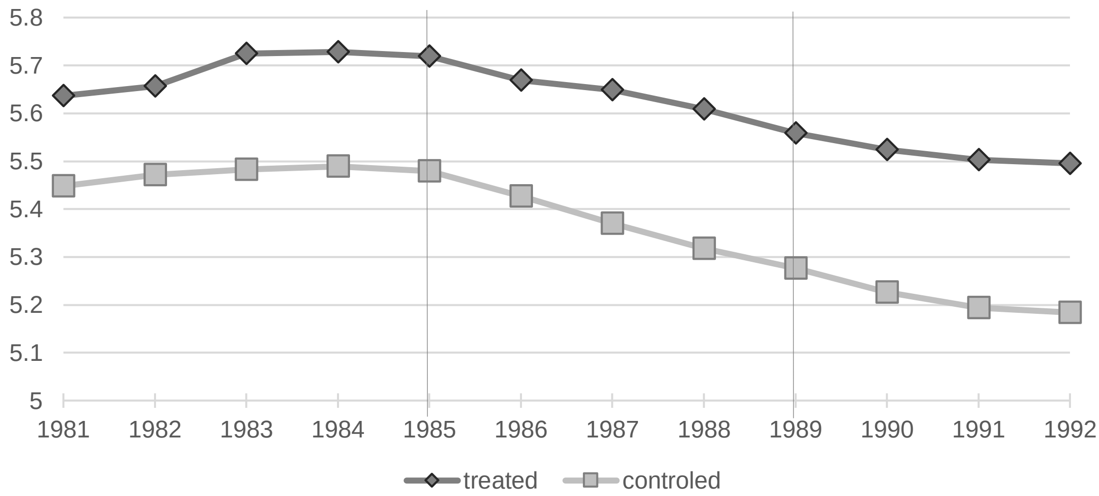
We also graph the analogous lines in Figure 3, using the sample we used for the estimation of (M5).[17] Since this sample now consists of only no-movers during FY1981–FY1992, these two lines may reveal the effects most noticeably if they existed at all. However, while the sizes of the treatment and control groups are smaller (n_k = 29 and n_l = 132), the trends of the two averages are similar to those in Figure 1. That is, after the two averages changed almost parallel from FY1983 to FY1986, the treated lifted slightly compared with the

[17] We also show an analogous figure for the PA expenses (inclusive of all costs for PA) and PA benefits (those that recipients actually receive) in Figures A1 and A2 in the Appendix. Their trends do not seem to be considerably different from those in Figure 3.

untreated, implying that the treated had a less negative difference for the 1985 change. The line for the treated becomes slightly flatter before FY1989, while the line for the untreated had a slightly larger dip after FY1989.

Of course, it is not obvious whether these differences are caused by chance or by changes in central funding. Therefore, in the next section, we estimate, rather than eyeball, the effects of the central funding using the models developed in Section 3, along with the data described in this section.

Figure 3 *Average PA caseloads (in logarithm) of the treated and controlled with no-movers for the entire period*



5. Estimation Results

5.1 Two-period DD Estimation

Tables 3 and 4 list the results of the DD estimation for the 1985 reduction and 1989 increase.[18] Each table contains the results of the 48 patterns of the DD estimation, with three types of ending points ($t_A = t_B + 1$, $t_B + 2$, and $t_B + 3$), four types of treatment variables (with or without movers, baseline, and longer horizon), and four types of DD estimators (baseline

[18] We obtain all the DD estimates by using a STATA module "diff" written by Villa (2016).

DD, DD with covariates, PSM-DD, and PSM-DD with common support). In the Appendix, we list the full results for (M2).

Table 3 *DD estimates for the 1985 matching rate reduction*

Periods	Samples	Treatment [#treated/#obs.]	Regression DD		Kernel propensity score matching DD		
			M1	M2	M3	M4 [#treated /#obs.]	
FY1984 −1985	Movers in the control group	Baseline [67/254]	0.007 (0.006)	0.006 (0.011)	0.008 (0.006)	0.007 (0.007)	[63/193]
		Longer horizon [30/254]	−0.005 (0.010)	0.011 (0.016)	0.023 (0.021)	0.017 (0.019)	[26/202]
	No movers	Baseline [67/241]	0.005 (0.007)	0.006 (0.013)	0.011 (0.008)	0.010 (0.010)	[62/185]
		Longer horizon [30/187]	−0.007 (0.010)	0.001 (0.020)	−0.000 (0.011)	−0.001 (0.013)	[18/112]
FY1984 −1986	Movers in the control group	Baseline [67/254]	0.011 (0.017)	−0.001 (0.017)	0.006 (0.014)	0.000 (0.014)	[63/193]
		Longer horizon [30/254]	−0.012 (0.019)	0.008 (0.025)	0.029 (0.042)	0.014 (0.037)	[26/202]
	No movers	Baseline [67/231]	0.008 (0.019)	−0.007 (0.021)	−0.000 (0.021)	0.001 (0.021)	[59/170]
		Longer horizon [30/187]	−0.013 (0.020)	−0.018 (0.027)	−0.013 (0.018)	−0.026 (0.022)	[18/112]
FY1984 −1987	Movers in the control group	Baseline [63/254]	0.034 (0.022)	0.020 (0.020)	0.018 (0.017)	0.014 (0.019)	[59/197]
		Longer horizon [30/254]	0.020 (0.022)	0.044 (0.025)	0.067 (0.041)	0.050 (0.034)	[26/202]
	No movers	Baseline [63/221]	0.031 (0.024)	0.010 (0.027)	−0.000 (0.023)	−0.014 (0.026)	[56/172]
		Longer horizon [30/187]	−0.013 (0.020)	0.022 (0.032)	0.016 (0.025)	−0.003 (0.035)	[18/112]

Notes: (i) ***: $p \leq 0.01$; **: $0.01 < p \leq 0.05$; *: $0.05 < p \leq 0.10$. No results are statistically significant above at these standard levels of significance. (ii) The standard errors between parentheses are based on clustering among cities in the same prefectures. (iii) "Baseline" refers to the treatment variable that takes unity when a city has not received a LAT grant in every FY between t_B and t_A. (iv) "Longer horizon" refers to the treatment variable that takes unity when a city has received a LAT grant in every FY between $t_B - 3$ and t_A. (v) "Movers" refer to cities whose LAT status was not constant either between t_B and t_A or between $t_B - 3$ and t_A. (vi) M1 refers to a DD regression without covariates; M2 to a DD regression with covariates; M3 to kernel propensity score matching DD without common support; and M4 to kernel propensity score matching DD with common support.

Table 3 lists the estimates of the 1985 matching rate reduction. If we side with the argument by the central government that central funding increases PA caseloads, we would expect cities without LAT receipts to have smaller PA caseloads than those with LAT receipts, as the former endured the full cost of the CGS-PA reduction, while the latter escaped it through the compensation made by the corresponding LAT increases. If this loosening effect were in place, we would expect the estimates to be *negative* and statistically significant. However, only one-third of the estimates have negative values (16 of 48), and in such cases, their sizes are small (as in other cases). More importantly, none of the estimates are statistically significant at the standard levels of significance. The smallest p value is 0.112 ((M2) for t_A = FY1987 with movers and longer-horizon treatment), which may be a borderline case. However, the p values for the other 47 cases are rather large, as shown by the histogram in Figure 4 that comprises the 48 p values from the estimates in Table 3.

Figure 4 *Histogram for the p values of the DD estimates for the 1985 change*

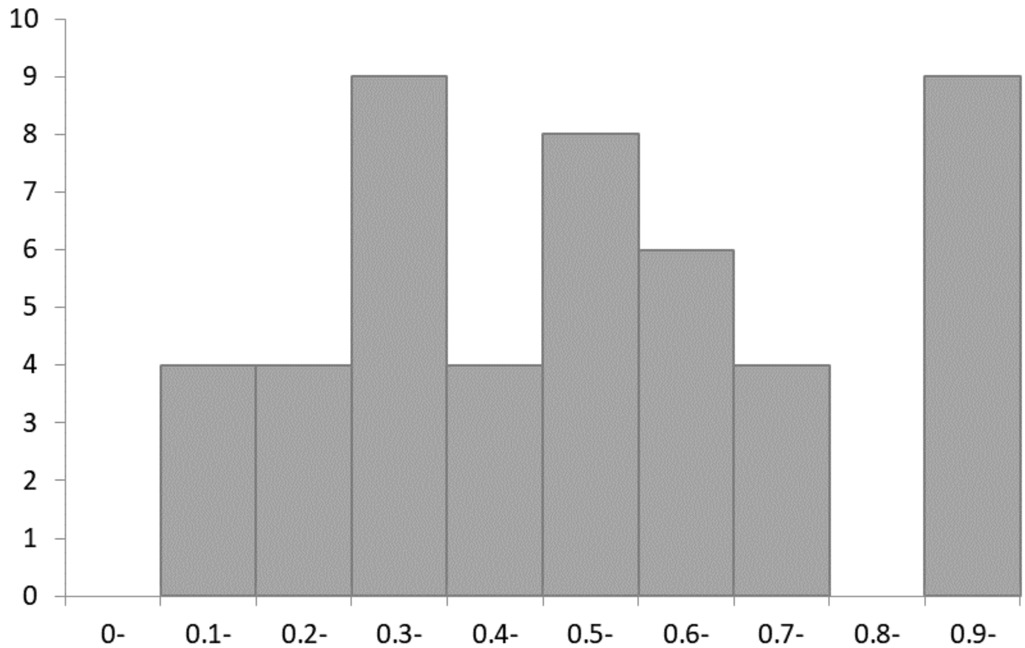


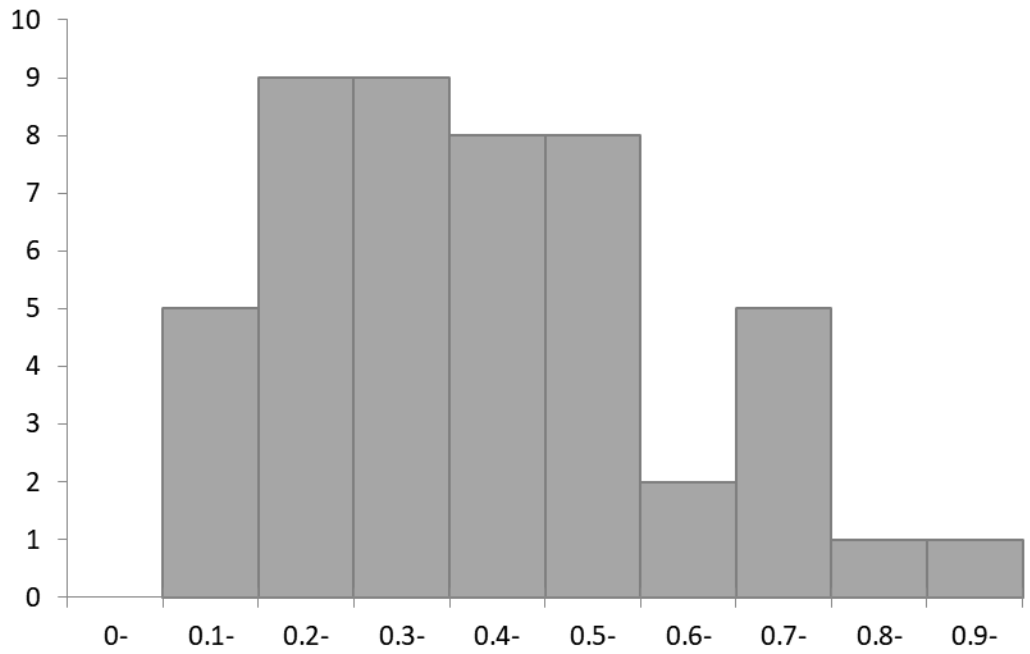
Table 4 *DD estimates for the 1989 matching rate increase*

Periods	Samples	Treatment [#treated/#obs.]	Regression DD		Kernel propensity score matching DD		
			M1	M2	M3	M4 [#treated /#obs.]	
FY1988 –1989	Movers in the control group	Baseline [90/254]	0.001 (0.009)	0.004 (0.011)	0.009 (0.016)	0.009 (0.015)	[78/218]
		Longer horizon [67/254]	0.005 (0.011)	0.011 (0.012)	0.006 (0.014)	0.005 (0.014)	[62/198]
	No movers	Baseline [90/232]	0.001 (0.009)	0.003 (0.011)	0.009 (0.013)	0.009 (0.013)	[82/200]
		Longer horizon [67/199]	0.004 (0.011)	0.011 (0.015)	0.015 (0.021)	0.018 (0.020)	[60/142]
FY1988 –1990	Movers in the control group	Baseline [89/254]	0.009 (0.011)	0.010 (0.014)	0.022 (0.020)	0.027 (0.017)	[76/220]
		Longer horizon [67/254]	0.019 (0.014)	0.021 (0.018)	0.022 (0.016)	0.014 (0.037)	[62/198]
	No movers	Baseline [89/228]	0.009 (0.011)	0.006 (0.015)	0.015 (0.017)	0.016 (0.016)	[59/170]
		Longer horizon [67/199]	0.017 (0.013)	0.014 (0.021)	0.032 (0.027)	0.038 (0.023)	[60/142]
FY1989 –1991	Movers in the control group	Baseline [87/254]	0.016 (0.019)	0.015 (0.018)	0.024 (0.027)	0.033 (0.024)	[75/217]
		Longer horizon [67/254]	0.028 (0.022)	0.024 (0.025)	0.034 (0.025)	0.023 (0.025)	[62/198]
	No movers	Baseline [87/222]	0.015 (0.019)	0.010 (0.016)	0.016 (0.023)	0.020 (0.022)	[79/193]
		Longer horizon [67/199]	0.027 (0.023)	0.016 (0.024)	0.045 (0.037)	0.054 (0.033)	[60/142]

Notes: (i) ***: $p \leq 0.01$; **: $0.01 < p \leq 0.05$; *: $0.05 < p \leq 0.10$. No results are statistically significant above at these standard levels of significance. (ii) The standard errors between parentheses are based on clustering among cities in the same prefectures. (iii) "Baseline" refers to the treatment variable that takes unity when a city has not received a LAT grant in every FY between t_B and t_A. (iv) "Longer horizon" refers to the treatment variable that takes unity when a city has received a LAT grant in every FY between $t_B - 3$ and t_A. (v) "Movers" refer to cities whose LAT status was not constant either between t_B and t_A or between $t_B - 3$ and t_A. (vi) M1 refers to a DD regression without covariates; M2 to a DD regression with covariates; M3 to kernel propensity score matching DD without common support; and M4 to kernel propensity score matching DD with common support.

Table 4 lists the 48 estimates for the 1989 matching rate increase. We now would expect cities without LAT receipts to increase their PA caseloads compared with those with LAT receipts, since the former this time would enjoy the full benefit of the CGS-PA increase, while the latter would miss out through the corresponding reduction in LAT disbursements. If this were the case, we would expect the estimate to be *positive* and statistically significant. The table shows that while all the estimates are positive, their sizes are small and none of them are statistically significant. The smallest p value among the 48 cases is 0.130 ((M4) for t_A = FY1991 without movers and with the longer-horizon treatments). Again, the p values for the other cases are rather large as shown in Figure 5 with the histogram of 48 p values obtained from the estimates in Table 4.

Figure 5 *Histogram for the p values of the DD estimates for the 1989 change*

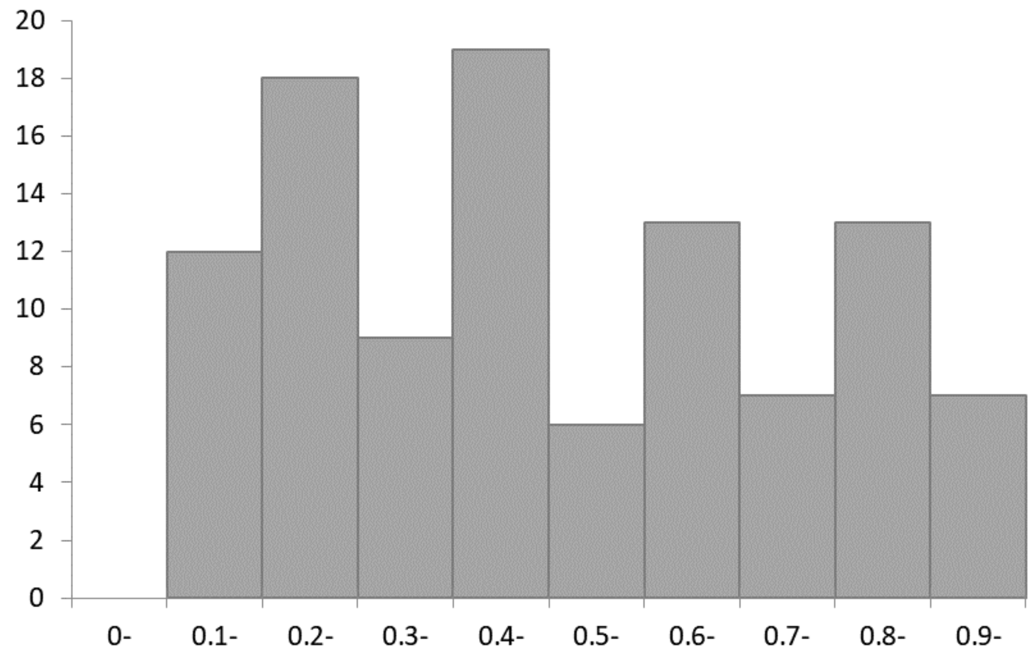


5.2 Joint Estimation for the 1985 and 1989 Changes

We estimate the various patterns of the generic model (M5) that jointly captures the effects of the 1985 and 1989 changes. There are 104 patterns, as the combinations of two types of time effect (year and prefecture-year), two types of time trend (linear and quadratic), four dynamic specifications (Q = 0, 1, 2, and 3), five estimation methods (the within estimator for the static model with Q = 0, and one-step difference, two-step difference, one-step system and two-step system GMM estimators for each of the dynamic models with Q = 1, 2, and 3), and two sorts of the sample (with movers or no-movers in the control group).

Table 5 lists the estimates of the effects of the 1985 matching rate reduction. Here, once again, we would expect the estimate to be negative and statistically significant. However, we obtain estimates with negative values for only about one-third of the cases (35 of 104), and in such cases, their values are small (as in the other cases). In addition, none of them is statistically significant at the standard levels. Meanwhile, Figure 6 exhibits a histogram that comprises the 104 p values from the estimates in Table 5. There are two possible borderline cases with p values of 0.108 and 0.117.[19] However, they obviously do not indicate the expected negative effects, as their coefficient values are all positive.

Figure 6 *Histogram for the p values of the estimates for the 1985 change with panel data 1981–1992*



[19] These two are the one-step difference estimates obtained from specifications with year effect and the first and second lags of the dependent variables, based on the sample that includes movers. The difference between the two is that the one with p = 0.108 is estimated with a quadratic trend, while the other with p = 0.117 is estimated with a linear trend.

Table 5 *Estimates for the 1985 matching rate reduction with 1981–1992 panel data*

Lags as regressors	Estimator		Movers in the control group		No movers		Movers in the control group		No movers	
			Liner time trend	Quadratic time trend	Liner time trend	Quadratic time trend	Liner time trend	Quadratic time trend	Liner time trend	Quadratic time trend
$Q = 0$	OLS (within)		–0.019	–0.029	–0.017	–0.028	–0.000	–0.021	–0.006	–0.023
			(0.022)	(0.020)	(0.023)	(0.021)	(0.021)	(0.020)	(0.023)	(0.020)
$Q = 1$: $\ln Y_{it-1}$	Difference estimator	One-step	–0.001	–0.010	–0.005	–0.016	0.003	–0.003	–0.005	–0.012
			(0.014)	(0.013)	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.017)
		Two-step	–0.006	–0.014	–0.005	–0.018	0.003	–0.005	-0.008	-0.017
			(0.015)	(0.016)	(0.014)	(0.016)	(0.017)	(0.018)	(0.020)	(0.025)
	System estimator	One-step	–0.003	–0.004	–0.004	–0.004	–0.008	–0.008	–0.013	–0.013
			(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.018)	(0.018)	(0.019)
		Two-step	–0.006	–0.008	0.000	–0.003	0.002	0.000	–0.012	–0.017
			(0.021)	(0.021)	(0.020)	(0.021)	(0.024)	(0.024)	(0.024)	(0.025)
$Q = 2$: $\ln Y_{it-1}$, $\ln Y_{it-2}$	Difference estimator	One-step	0.020	0.024	0.009	0.004	0.028	0.037	0.024	0.036
			(0.018)	(0.020)	(0.019)	(0.022)	(0.018)	(0.023)	(0.019)	(0.025)
		Two-step	0.018	0.021	0.012	0.007	0.027	0.036	0.006	0.025
			(0.019)	(0.031)	(0.062)	(0.025)	(0.020)	(0.026)	(0.023)	(0.031)
	System estimator	One-step	0.018	0.018	0.015	0.015	0.012	0.013	0.009	0.009
			(0.018)	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)
		Two-step	0.017	0.015	0.027	0.024	0.015	0.015	0.011	0.010
			(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.021)	(0.023)
$Q = 3$: $\ln Y_{it-1}$, $\ln Y_{it-2}$, $\ln Y_{it-3}$	Difference estimator	One-step	0.012	–0.001	0.009	–0.013	0.020	0.015	0.021	0.010
			(0.017)	(0.018)	(0.019)	(0.020)	(0.017)	(0.020)	(0.020)	(0.023)
		Two-step	0.012	–0.001	0.011	–0.015	0.024	0.019	0.015	0.006
			(0.025)	(0.019)	(0.020)	(0.023)	(0.019)	(0.023)	(0.021)	(0.025)
	Systems estimator	One-step	0.022	0.028	0.022	0.025	0.019	0.026	0.019	0.027
			(0.020)	(0.020)	(0.021)	(0.021)	(0.020)	(0.020)	(0.021)	(0.023)
		Two-step	0.020	0.024	0.031	0.034	0.019	0.026	0.017	0.025
			(0.022)	(0.023)	(0.021)	(0.023)	(0.022)	(0.024)	(0.023)	(0.026)
Year effects or prefecture-year effects			Year	Year	Year	Year	Pref-year	Pref-year	Pref-year	Pref-year

Notes: (i) ***: $p \leq 0.01$; **: $0.01 < p \leq 0.05$; *: $0.05 < p \leq 0.10$. No results are statistically significant above at these standard levels of significance. (ii) The standard errors between parentheses are based on clustering along the time dimension. (iii) "Movers" refer to cities whose LAT status was not constant between 1981 and 1992. (iv) "Year effects (Year)" refer to cases where a regression model includes year dummies. (v) "Prefecture-year effects (Pref-year)" refer to cases where a regression model includes interaction terms between prefecture and year dummies. (vi) There are 254 and 161 cross section units in the sample for the cases with and without movers, respectively. (vii) The number of periods in the sample are 12 for the within estimates; 11, 10, and 9 for the system estimates with $Q = 1$, 2, and 3, respectively; and 10, 9, and 8 for the difference estimates with $Q = 1$, 2, and 3, respectively.

Table 6 lists the estimates of the effects of the 1989 matching rate increase. For the 1989 change, we would expect the estimates to be *positive* and statistically significant. However, only less than one-fourth (24 of 104) are positive, and in such cases, their values are very small (as in the other cases). In addition, none of the positive estimates is statistically significant.

Three of the 104 estimates are statistically significant with p = 0.07, 0.050, and 0.092. However, their values are all negative, implying effects in the opposite direction to popular prediction. In addition, diagnostics imply misspecifications. Two of the three cases (with p = 0.07 and 0.050) use the static specifications with prefecture-year effect and linear time trend.[20] Meanwhile, we also estimate dynamic specifications that nest the two static specifications for Table 6. Since we find that their coefficients on the first lag of the dependent variable (ρ_1) are all estimated significantly with virtual zero p values, we reject the static specifications (cases *with* significant effect) in favor of the dynamic specifications (cases *without* significant effect).[21] The third case (with p = 0.092) comes from the one-step system estimation with year effect, liner time trend and the first lag of the dependent variable, and uses the sample without movers in the control group. However, the Arellano–Bond test suggests the existence of the AR(2) error, while the tests of over-identifying restrictions emphatically reject the null hypothesis, which shows that the effect is poorly estimated. In addition, there is one possible borderline case with p = 0.104.[22] However, its coefficient estimate has a *negative* value, although we would expect it to be positive.

---

[20] The difference between the two is that one with p = 0.070 uses the sample with movers, while the other with p = 0.050 uses the sample without movers.

[21] Details are not listed but are available on request.

[22] One-step system estimate obtained from the specification with prefecture-year effect and the first lag of the dependent variable, based on the sample that includes movers.

Table 6 *Estimates for the 1989 matching rate increase with 1981–1992 panel data*

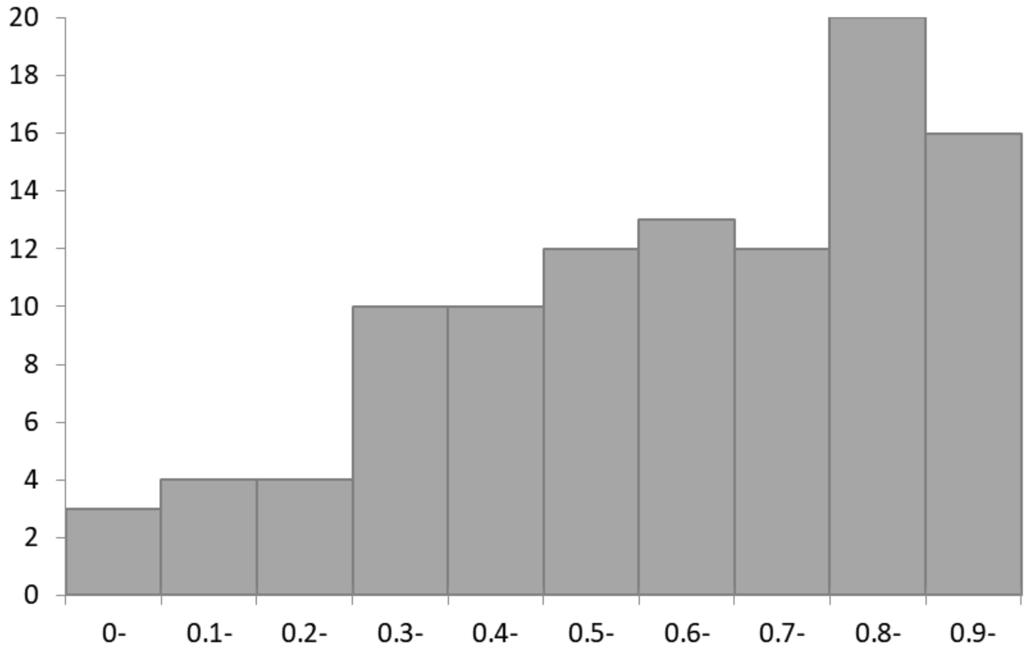
Lags as regressors	Estimator		Movers in the control group		No movers		Movers in the control group		No movers	
			Liner time trend	Quadratic time trend	Liner time trend	Quadratic time trend	Liner time trend	Quadratic time trend	Liner time trend	Quadratic time trend
No.	OLS (within)		–0.027	–0.017	–0.029	–0.018	–0.047*	–0.026	–0.055**	–0.026
			(0.027)	(0.021)	(0.028)	(0.022)	(0.026)	(0.020)	(0.028)	(0.019)
$\ln Y_{it-1}$	Difference estimator	One-step	–0.020	–0.016	–0.016	–0.016	–0.020	–0.018	–0.009	–0.005
			(0.021)	(0.020)	(0.021)	(0.014)	(0.020)	(0.019)	(0.020)	(0.020)
		Two-step	–0.007	–0.003	–0.007	–0.004	–0.008	–0.006	0.002	0.005
			(0.015)	(0.015)	(0.015)	(0.014)	(0.017)	(0.017)	(0.027)	(0.023)
	System estimator	One-step	–0.022	–0.016	–0.026*	–0.016	–0.026	–0.021	–0.022	–0.018
			(0.015)	(0.018)	(0.015)	(0.018)	(0.016)	(0.018)	(0.018)	(0.019)
		Two-step	–0.020	–0.008	–0.019	–0.008	–0.021	–0.014	–0.012	–0.007
			(0.016)	(0.016)	(0.016)	(0.017)	(0.021)	(0.021)	(0.023)	(0.023)
$\ln Y_{it-1}$, $\ln Y_{it-2}$	Difference estimator	One-step	–0.005	–0.007	–0.004	–0.005	-0.003	-0.006	0.015	0.010
			(0.021)	(0.021)	(0.021)	(0.021)	(0.020)	(0.020)	(0.019)	(0.019)
		Two-step	0.006	0.004	0.003	0.001	0.005	0.004	0.013	0.013
			(0.016)	(0.030)	(0.020)	(0.016)	(0.019)	(0.018)	(0.022)	(0.021)
	System estimator	One-step	–0.007	–0.005	–0.013	–0.005	–0.009	–0.010	–0.006	–0.003
			(0.015)	(0.018)	(0.015)	(0.018)	(0.015)	(0.018)	(0.017)	(0.019)
		Two-step	–0.002	0.003	–0.004	0.005	–0.002	–0.002	0.000	0.004
			(0.014)	(0.017)	(0.015)	(0.017)	(0.016)	(0.018)	(0.017)	(0.019)
$\ln Y_{it-1}$, $\ln Y_{it-2}$, $\ln Y_{it-3}$	Difference estimator	One-step	–0.016	–0.016	–0.016	–0.016	–0.013	–0.013	–0.001	–0.001
			(0.021)	(0.021)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
		Two-step	–0.003	–0.004	–0.007	–0.009	–0.004	–0.004	0.002	0.002
			(0.026)	(0.015)	(0.016)	(0.016)	(0.018)	(0.018)	(0.020)	(0.020)
	Systems estimator	One-step	–0.003	–0.011	–0.009	–0.011	–0.004	–0.014	0.002	–0.008
			(0.016)	(0.018)	(0.016)	(0.019)	(0.016)	(0.019)	(0.017)	(0.020)
		Two-step	0.001	–0.002	0.000	–0.001	0.003	-0.004	0.007	0.002
			(0.014)	(0.016)	(0.014)	(0.016)	(0.016)	(0.018)	(0.018)	(0.019)
Year effects or prefecture-year effects			Year	Year	Year	Year	Pref-year	Pref-year	Pref-year	Pref-year

Notes: (i) ***: $p \leq 0.01$; **: $0.01 < p \leq 0.05$; *: $0.05 < p \leq 0.10$. No results are statistically significant above at these standard levels of significance. (ii) The standard errors between parentheses are based on clustering along the time dimension. (iii) "Movers" refer to cities whose LAT status was not constant either between 1981 and 1992. (iv) "Year effects (Year)" refer to cases where a regression model includes year dummies. (v) "Prefecture-year effects (Pref-year)" refer to cases where a regression model includes interaction terms between prefecture and year dummies. (vi) There are 254 and 161 cross section units in the sample for the cases with and without movers, respectively. (vii) The number of periods in the sample are 12 for the within estimates; 11, 10, and 9 for the system estimates with Q = 1, 2, and 3, respectively; and 10, 9, and 8 for the difference estimates with Q = 1, 2, and 3, respectively.

Figure 7 exhibits a histogram that comprises the 104 *p* values from the estimates in Table 6. The distribution of *p* values is more negatively skewed in Figure 6 than in Figure 5. In other words, despite the three significant cases and one borderline case, the estimates of the 1989 effect tend to have larger *p* values than those of the 1985 effect.

Therefore, while the sets of results in this section suggest that despite the three cases of statistically significant estimates in Table 6 and the four cases of borderline estimates in Tables 4, 5 and 6, our analysis cannot substantiate the existence of the loosening effect of central funding on local welfare caseloads.

Figure 7 *Histogram for the p values of the estimates for the 1989 change with panel data 1981–1992*



6. Concluding Remarks

This study examined the loosening effects of central funding on the size of local welfare caseloads. We took advantage of an institutional mechanism of the Japanese system of central grants and two historical changes in the matching rate for local PA programs. By performing various estimations, we argued that neither the 1985 nor the 1989 changes affected PA caseloads, as the loosening effect suggests.

When estimation yields a result that is not statistically significant, we may be inclined to downplay it, as we can easily come up with multiple explanations (Hewitt et al., 2008). For example, we could argue that a non-significant result is due to the lack of power, using terms such as "borderline" significance and claiming that the effect might in fact exist. We could also underemphasize the result by simply stating that no firm conclusions can be drawn, as the sample size is insufficiently large. Furthermore, we could blame the inappropriate use of a given estimator that may have produced a non-significant result. Finally, such a tendency to downplay might be more noticeable, especially if the result conflicts with a well-known assumption in public policy with vested interests involved.

Our results may be robust to these reservations. Of the 256 ($= 48 \times 2 + 104 \times 2$) estimates in Tables 3–6, there were only three statistically significant estimates. Moreover, these estimates had negative signs, being opposite to what the loosening effect suggests, and our diagnostics suggested misspecifications. There were four borderline estimates among the remaining 253 cases that were statistically insignificant. However, all four had opposite signs to what the loosening effect implies.

However, our results may have shortcomings originating in our sample choice. First, the size of our cross-section units was modest (N = 254).[23] We could have nonetheless increased the size by including the data from cities in non-urban prefectures. However, such an inclusion might make the sample more unbalanced. In addition, since LAT localities are concentrated in non-urban areas, it will only increase the relative size

[23] However, this is substantially larger than those in US studies of welfare caseloads. The typical size of the cross-section units is the number of US states (i.e., 51). Other studies use county data (Schiller and Basher, 1993) or regional labor market data (Page et al., 2005), with their sizes less than 100.

of the control group. In this regard, the literature indicates that increasing only the control group size may not only fail to improve the estimation but also invalidate the standard procedures for inferences (Conley and Taber, 2011; MacKinnon and Webb, 2016). Second, being based on a sample of cities in urban prefectures, our results may lack external validity. They may not be applicable when we consider the effect of central funding on cities in non-urban prefectures. Our results are nonetheless relevant to the policy dialog on PA programs in Japan, since the central government typically refers to specific cities in "urban" prefectures when it maintains the loosening effect of central funding (Kimura 2006). In summary, we regard our choice of a sample as the best compromise given the limited availability of PA caseload data, regional distribution of LAT localities, and relevance to specific policy issues in Japan.

Our failure to substantiate the loosening effect may point to the dominance of central control over local discretion and suggest that administrative factors rather than fiscal ones are the most important. Therefore, while it would be hard to obtain measures that index a variety of central administrative controls over localities, it is indeed important to empirically examine the effects of such factors on PA caseloads. The next step would thus be to explore such intergovernmental administrative aspects of the PA system.

References

Andini, C. (2006), Unemployment and Welfare Participation in a Structural VAR: Rethinking the 1990s in the United States, International Review of Applied Economics 20, 243–253.

Angrist, J. D., Pischke, S. (2009), Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press.

Arellano, M., Bond, S. (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277–297.

Arellano, M., Bover, O. (1995), Another look at the instrumental variable estimation of error-components models. Journal of Econometrics 68, 29–51.

Autor, D. (2003), Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing, Journal of Labor Economics 21, 1–42.

Ayala, L., Pérez, C. (2005), Macroeconomic Conditions, Institutional Factors and Demographic Structure: What Causes Welfare Caseloads? Journal of Population Economics 18, 563–581.

Ayala, L., Cano, Á.T. (2017), Economic Downturns, Endogenous Government Policy and Welfare Caseloads, Hacienda Pública Española/Review of Public Economics 220, 107–136.

Baicker, K. (2005), Extensive or Intensive Generosity? The Price and Income Effects of Federal Grants, Review of Economics and Statistics 87, 371–384.

Baker, M., Payne, A.A., Smart, M. (1998), An Empirical Study of Matching Grants: The "Cap on CAP," Journal of Public Economics 72, 269–288.

Berg, N., Gabel, T. (2015), Did Canadian Welfare Reform Work? The Effects of New Reform Strategies on Social Assistance Participation, Canadian Journal of Economics 48, 498–528.

Besley, T., Burgess, R. (2004), Can Labor Regulation Hinder Economic Performance? Evidence from India, Quarterly Journal of Economics 119, 91–134.

Blank, R.M. (2001), What Causes Public Assistance Caseloads to Grow? Journal of Human Resources 36, 85–118.

Blundell, R., Bond, S. (1998), Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. Journal of Econometrics 87, 115–143.

Buettner, T. (2006), The Incentive Effect of Fiscal Equalization Transfers on Tax Policy, Journal of Public Economics 90, 477–497.

Cadena, B., Danziger, S., Seefeldt, K. (2006), Measuring State Welfare Policy Changes: Why Don't They Explain Caseload and Employment Outcomes? Social Science Quarterly 87, 808–817.

Cattaneo, M.D., Jansson, M., Ma, X. (2017a), Simple Local Polynomial Density Estimators, Working Paper, University of Michigan.

Cattaneo, M.D., Jansson, M., Ma, X. (2018), Manipulation Testing based on Density Discontinuity, Stata Journal 18, 234–261.

Chernick, H. (1998), Fiscal Effects of Block Grants for the Needy: An Interpretation of the Evidence, International Tax and Public Finance 5, 205–233.

Conley, T.G., Taber, C.R. (2011), Inference with "Difference in Differences" with a Small Number of Policy Changes, Review of Economics and Statistics 93, 113–125.

Dahlberg, M., Mörk, E., Rattsø, J., Ågren, H. (2008), Using a Discontinuous Grant Rule to Identify the Effect of Grants on Local Taxes and Spending, Journal of Public Economics 92, 2320–2335.

Danielson, C., Klerman, J.A. (2008), Did Welfare Reform Cause the Caseload Decline? Social Service Review 82, 703–730.

Deaton, A., Muellbauer, J. (1980), Economic and Consumer Behavior. Cambridge University Press.

Friedberg, L. (1998), Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data, American Economic Review 88, 608–627.

Gordon, N. (2004), Do Federal Grants Boost School Spending? Evidence from Title I, Journal of Public Economics 88, 1771–1792.

Grubb, N.W. (1984), The Price of Local Discretion: Inequalities in Welfare Spending within Texas, Journal of Policy Analysis and Management 3, 359–372.

Gustafsson, B. (1984), Macroeconomic Performance, Old Age Security and the Rate of Social Assistance Recipients in Sweden, European Economic Review 26, 319–338.

Heckman, J.J., Ichimura, H., Todd, P.E. (1997), Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, Review of Economic Studies 64, 605–654.

Hewitt, G.H., Mitchell, N., Torgerson, D. (2008), Heed the Data when Results are not Significant, BMJ 336, 23–25.

Hill, B.C., Murray, M.N. (2008), Interactions between Welfare Caseloads and Local Labor Markets, Contemporary Economic Policy 26, 539–554.

Holtz-Eakin, D. (1986), Unobserved Tastes and the Determination of Municipal Services, National Tax Journal 39, 527–532.

Holtz-Eakin, D., Newey, W., and Rosen, H.S. (1988), Estimating Vector Autoregressions with Panel Data, Econometrica 56, 1371–1395.

Huang, C., Garfinkel, I., Waldfogel, J. (2004), Child Support Enforcement and Welfare Caseloads, Journal of Human Resources 39, 108–134.

Jagannathan, R. (2011), Welfare Reform’s Impact on Caseload Decline in the United States: An Application Latent Trajectory Model, Social Science Journal 48, 703–721.

Johnson, T.R., Klepinger, D.H., Dong, F.B. (1994), Caseload Impacts of Welfare Reform, Contemporary Economic Policy 12, 89–101.

Kasai, Y. (1978), Interview with Yoshisuke Kasai, transcript. in Zaidanhojin Shakai Fukushi Kenkyujo (Japanese Research Institute on Social Welfare, Inc.) Ed. Senryoki ni okeru Shakai Fukushi Shiryo ni kansuru Kenkyu Hokokusho (Report on Studies of Social Welfare Documents and Data during the Occupation of Japan), 278–307.

Keiser, L.R., Soss, J. (1998), With Good Cause: Bureaucratic Discretion and the Politics of Child Support Enforcement, American Journal of Political Science 42, 1133–1156.

Klerman, J.A., Danielson, C. (2016), Can the Economy Explain the Explosion in the Supplemental Nutrition Assistance Program Caseload? An Assessment of the Local-level Approach, American Journal of Agricultural Economics 98, 92–112.

Kobayashi, H. (2014), The Future of the Public Assistance Reform in Japan: Workfare versus Basic Income? in: Vanderborght, Y., and Yamamori, T. (Eds.), Basic Income in Japan Prospects for a Radical Idea in a Transforming Welfare State, Springer, Berlin, 83–99.

Kneebone, R.D., While, K.G. (2009), Fiscal Retrenchment and Social Assistance in Canada, Canadian Public Policy 35, 21–40.

Kimura, Y. (2006), Seikatsu Hogo no kyogikai ni kakawatte (Participating in the Consultative Meeting on Public Assistance and Child Rearing Allowance), Chiho Zaisei, March 2006.

Leuven, E., Sianesi, B. (2003), psmatch2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. Statistical Software Components S432001, Department of Economics, Boston College.

MacKinnon, J.G., Webb, M.D. (2016), Randomization Inference for Difference-in-differences with Few Treated Clusters. Working Paper 1355, Queen’s University, Department of Economics.

Nakai, H. (1988), Gendai Zaiseifutan no Suryo Bunseki (Modern Fiscal Burden: A Quantitative Analysis), Yuhikaku, Tokyo.

Nerlove, M. (1956), Estimates of the Elasticities of Supply of selected Agricultural Commodities. Journal of Farm Economics 38, 496–509.

Nickell, S. (1981), Biases in Dynamic Models with Fixed Effects. Econometrica 49, 1417–1426.

Okuno, M. (1944), Shi-cho-son zaisei no jittai to kokuhi chiho futan kubun (2) (Municipal Finances and Fiscal Sharing of National and Local Expenditures: Part 2), Jichi Kenkyu 20, 15–20.

Page, M.E., Spetz, J., Millar, J. (2005), Does the Minimum Wage Affect Welfare Caseloads? Journal of Policy Analysis and Management 24, 273–295.

Ribar, D.C., Wilhelm, M.O. (1999), The Demand for Welfare Generosity, Review of Economics and Statistics 81, 96–108.

Roodman, D. M. (2009a), How to Do xtabond2: An Introduction to Difference and System GMM in Stata. Stata Journal 9, 86–136.

Roodman, D. M. (2009b), A Note on the Theme of Too Many Instruments. Oxford Bulletin of Economics and Statistics 71, 135–158.

Schiller, B.R. (1999), State Welfare-reform Impacts: Content and Enforcement Effects, Contemporary Economic Policy 17, 210–222.

Schiller, B.R., Brasher, C.N. (1993), Effects of Workfare Saturation on AFDC Caseloads, Contemporary Economic Policy 11, 39–49.

Snarr, H.W. (2011), Was it the Economy or Reform that Precipitated the Steep Decline in the US Welfare Caseload? Applied Economics 45, 525–540.

Spindler, Z.A., Gilbreath, W.S. (1979), Determinants of Canadian Social Assistance Participation Rates, International Journal of Social Economics 6, 164–176.

Suzuki, W., Zhou, Y. (2007), Welfare Use in Japan: Trends and Determinants, Journal of Income Distribution 16, 88–109.

Villa, J.M. (2016), diff: Simplifying the Estimation of Difference-in-differences Treatment Effects, Stata Journal 16, 52–71.

Weissert, C.S. (1994), Beyond the Organization: The Influence of Community and Personal Values on Street-level Bureaucrats' Responsiveness, Journal of Public Administration Research and Theory 4, 225–254.

Windmeijer, F. (2005), A Finite Sample Correction for the Variance of Linear Efficient Two-step GMM Estimators. Journal of Econometrics 126, 25–51.

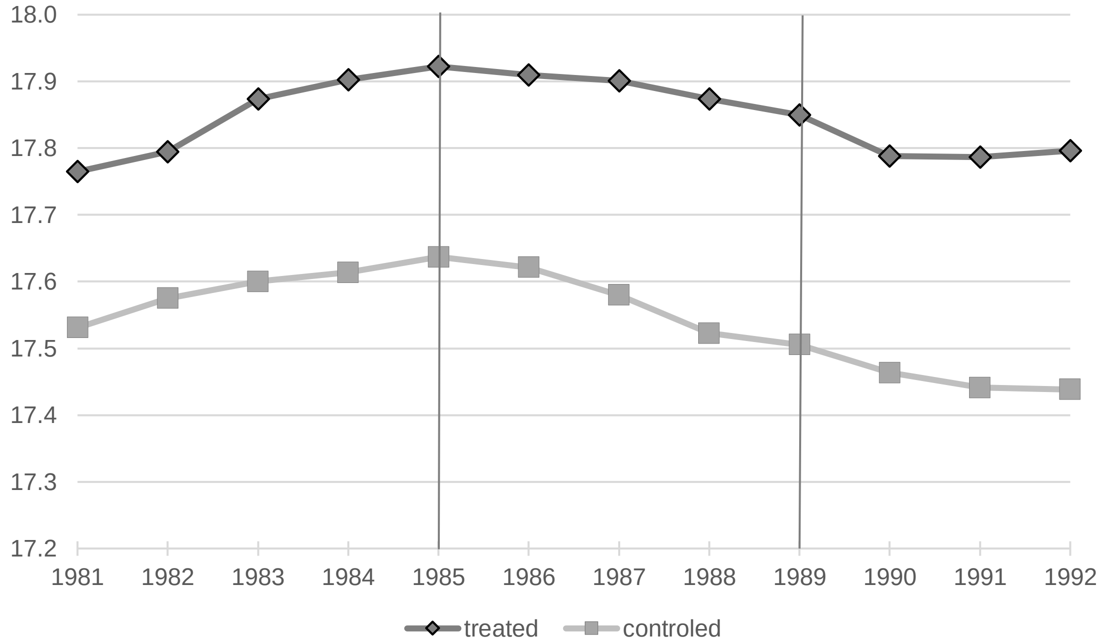
Woodridge, J.M., (2010), Econometric Analysis of Cross Section and Panel Data, MIT Press.

Wolfers, J. (2006), Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results, American Economic Review, 96(5), pp. 1802-1820. Review, 88(3), pp. 608-627.

Ziliak, J.P., Figlio, D.N., Davis, E.E., Connolly, L.C. (2000), Accounting for the Decline in AFDC Caseloads: Welfare Reform or the Economy? Journal of Human Resources 35, 570–586.

Appendix

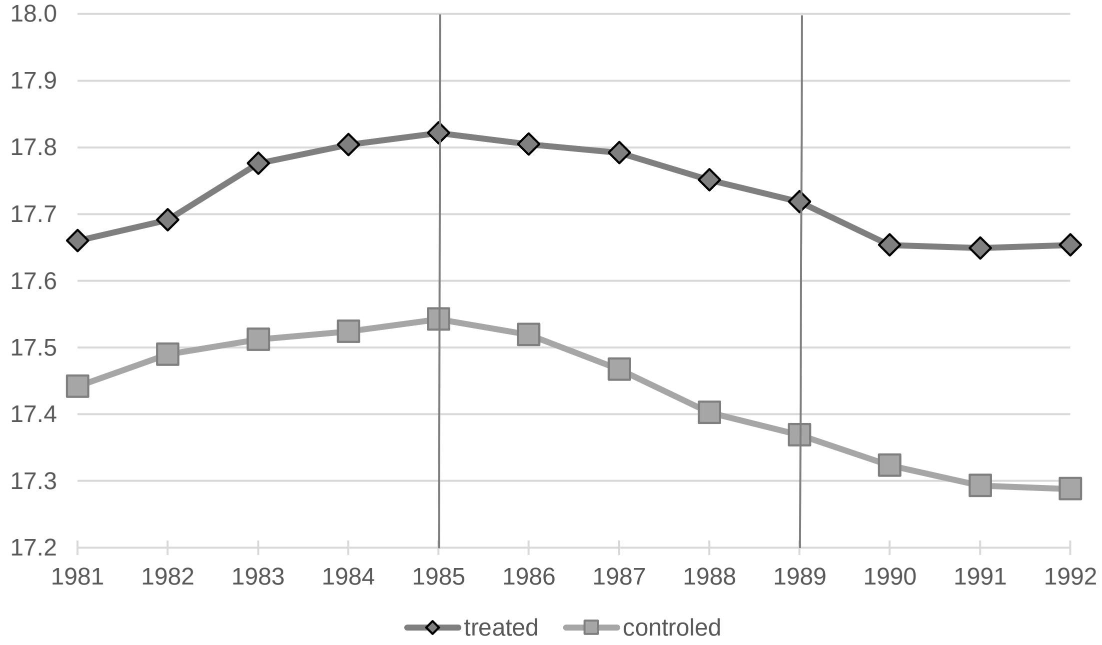
Figure A1 *Average PA expenditures (in logarithm) of the treated and controlled with no-movers for the entire period*



Notes: The data for PA expenditures are obtained from *Chihozaisei Jokyo Chosa* (Annual Survey on Local Public Finance), compiled by the Ministry of Internal Affairs and Communications, which is available at e-Stat (https://www.e-stat.go.jp).

Figure A2

Average PA payments (in logarithm) of the treated and controlled with no-movers for the entire period



Notes: The data for PA payments are obtained from *Chihozaisei Jokyo Chosa* (Annual Survey on Local Public Finance), compiled by the Ministry of Internal Affairs and Communications, which is available at e-Stat (https://www.e-stat.go.jp).

Table A1 *Full results for Model M2 for the 1985 matching rate reduction*

	FY1984–1985				FY1984–1986				FY1984–1987			
	Movers in the control group		No movers		Movers in the control group		No movers		Movers in the control group		No movers	
	Baseline	Longer horizon	Baseline	Longer horizon	Baseline	Longer horizon	Baseline	Longer horizon	Baseline	Longer horizon	Baseline	Longer horizon
Treatment	0.006	0.011	0.006	0.001	–0.001	0.008	–0.007	–0.018	0.020	0.044	0.010	0.022
	(0.011)	(0.016)	(0.013)	(0.020)	(0.017)	(0.025)	(0.020)	(0.027)	(0.020)	(0.025)	(0.027)	(0.032)
Poverty index	0.368*	0.377*	0.407*	0.367	0.269	0.299	0.208	0.119	0.100	0.156	0.089	–0.044
	(0.182)	(0.183)	(0.209)	(0.258)	(0.219)	(0.221)	(0.271)	(0.363)	(0.317)	(0.303)	(0.363)	(0.446)
Unemployment rate	0.688	0.719	0.082	0.494	2.973**	3.037*	2.413	3.095	2.981*	3.058*	1.959	2.131
	(0.652)	(0.726)	(0.572)	(0.790)	(1.229)	(1.382)	(1.357)	(1.864)	(1.384)	(1.420)	(1.850)	(1.990)
ln(population)	–0.002	–0.002	–0.003	–0.006	–0.002	–0.002	–0.003	–0.007	–0.002	–0.001	–0.003	–0.007
	(0.006)	(0.006)	(0.006)	(0.006)	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.008)	(0.012)
ln(households)	0.021	0.021	0.035	–0.019	–0.011	0.001	–0.047	–0.154	–0.034	–0.024	–0.091	–0.184
	(0.076)	(0.072)	(0.085)	(0.107)	(0.109)	(0.109)	(0.122)	(0.143)	(0.087)	(0.097)	(0.093)	(0.119)
Share of Korean residents	0.579	0.583	0.783	1.094	1.747	1.763	1.844	2.204	3.554*	3.542*	3.789	3.827
	(0.949)	(0.939)	(1.010)	(1.046)	(1.756)	(1.763)	(1.993)	(2.472)	(1.685)	(1.721)	(2.330)	(2.709)
Share of single mother households	–1.490	–1.454	–0.705	–2.595	–2.180	–2.195	–0.912	–3.818	–0.025	0.072	–1.315	–1.952
	(2.132)	(2.094)	(2.242)	(2.619)	(3.289)	(3.279)	(3.736)	(4.338)	(4.186)	(4.067)	(4.125)	(5.097)
Share of elderly households	–0.790**	–0.818**	–0.857**	–1.134***	–1.979***	–2.010***	–2.140***	–2.553***	–1.499*	–1.623**	–1.767*	–2.044*
	(0.298)	(0.263)	(0.293)	(0.350)	(0.515)	(0.485)	(0.557)	(0.655)	(0.687)	(0.678)	(0.814)	(0.961)
Share of workers in the manufacturing sector	0.121	0.121	0.144	0.086	0.004	0.012	–0.071	–0.166	–0.198	–0.196	–0.282	–0.431
	(0.143)	(0.140)	(0.149)	(0.170)	(0.127)	(0.113)	(0.154)	(0.179)	(0.230)	(0.222)	(0.283)	(0.283)
Share of workers in the service sector	0.428**	0.434**	0.497**	0.427	0.466*	0.485**	0.412	0.264	0.382	0.419	0.294	0.105
	(0.184)	(0.185)	(0.216)	(0.246)	(0.218)	(0.210)	(0.285)	(0.308)	(0.297)	(0.286)	(0.383)	(0.371)
Constant	–0.370	–0.378	–0.428	–0.230	–0.360	–0.395	–0.226	0.135	–0.267	–0.327	–0.096	0.265
	(0.266)	(0.269)	(0.293)	(0.365)	(0.300)	(0.310)	(0.346)	(0.401)	(0.341)	(0.340)	(0.404)	(0.397)
Sample size	254	254	241	187	254	254	231	187	254	254	221	187
R^2	0.177	0.178	0.195	0.188	0.222	0.223	0.228	0.240	0.217	0.222	0.234	0.227

Notes: (i) ***: $p \leq 0.01$; **: $0.01 < p \leq 0.05$; *: $0.05 < p \leq 0.10$. (ii) The standard errors between parentheses are based on clustering among cities in the same prefectures. (iii) "Baseline" refers to the treatment variable that takes unity when a city has not received a LAT grant in every FY between t_B and t_A. (iv) "Longer horizon" refers to the treatment variable that takes unity when a city has received a LAT grant in every FY between $t_B - 3$ and t_A. (v) "Movers" refer to cities whose LAT status was not constant either between t_B and t_A or between $t_B - 3$ and t_A.

Table A2 *Full results for Model M2 for the 1989 matching rate increase*

	FY1989–1990				FY1989–1991				FY1989–1992			
	Movers in the control group		No movers		Movers in the control group		No movers		Movers in the control group		No movers	
	Baseline	Longer horizon	Baseline	Longer horizon	Baseline	Longer horizon	Baseline	Longer horizon	Baseline	Longer horizon	Baseline	Longer horizon
Treatment	0.004	0.011	0.003	0.011	0.010	0.021	0.006	0.014	0.015	0.024	0.010	0.016
	(0.011)	(0.012)	(0.011)	(0.015)	(0.014)	(0.018)	(0.015)	(0.021)	(0.018)	(0.025)	(0.016)	(0.024)
Poverty index	0.160	0.198	0.101	0.185	–0.015	0.049	–0.118	0.076	-0.306	–0.244	–0.415	–0.239
	(0.190)	(0.186)	(0.196)	(0.186)	(0.322)	(0.325)	(0.340)	(0.302)	(0.438)	(0.425)	(0.426)	(0.417)
Unemployment rate	0.279	0.381	0.471	0.393	0.000	0.155	0.185	–0.486	–0.216	–0.067	–0.273	–0.681
	(0.951)	(0.960)	(0.941)	(0.940)	(1.582)	(1.608)	(1.603)	(1.225)	(1.645)	(1.722)	(1.792)	(1.360)
ln(population)	–0.002	–0.002	–0.001	0.001	0.009	0.008	0.010	0.014	0.018	0.017	0.017	0.026
	(0.005)	(0.005)	(0.006)	(0.007)	(0.009)	(0.010)	(0.011)	(0.014)	(0.012)	(0.013)	(0.015)	(0.019)
ln(households)	–0.098	–0.087	–0.105	–0.083	–0.141	–0.126	–0.164	–0.148	0.013	0.021	0.018	0.015
	(0.061)	(0.054)	(0.064)	(0.074)	(0.088)	(0.087)	(0.087)	(0.118)	(0.100)	(0.099)	(0.093)	(0.101)
Share of Korean residents	–0.173	–0.155	–0.523	–0.937	–1.361	–1.331	–2.522	–3.694	–2.926	–2.911	–5.355	–7.517
	(0.834)	(0.816)	(1.155)	(1.173)	(1.543)	(1.541)	(1.759)	(1.873)	(1.639)	(1.616)	(3.185)	(3.282)
Share of single mother households	1.296	1.212	1.653	2.327	3.032	2.863	3.882	4.471	3.245	3.062	4.225	4.276
	(1.779)	(1.712)	(1.966)	(1.979)	(2.661)	(2.573)	(2.953)	(2.155)	(3.626)	(3.535)	(3.540)	(2.839)
Share of elderly households	–0.278	–0.311	–0.270	–0.348	0.086	0.018	0.065	–0.202	0.509	0.415	0.511	0.373
	(0.579)	(0.526)	(0.607)	(0.625)	(0.702)	(0.694)	(0.785)	(0.827)	(1.091)	(1.083)	(1.200)	(1.211)
Share of workers in the manufacturing sector	–0.040	–0.027	–0.082	–0.047	–0.156	–0.132	–0.217	–0.140	–0.091	–0.072	–0.080	–0.055
	(0.161)	(0.159)	(0.157)	(0.174)	(0.215)	(0.212)	(0.231)	(0.244)	(0.311)	(0.300)	(0.346)	(0.351)
Share of workers in the service sector	–0.092	–0.072	–0.145	–0.112	–0.322	–0.287	–0.406	–0.269	–0.280	–0.248	–0.263	–0.187
	(0.180)	(0.177)	(0.181)	(0.201)	(0.285)	(0.287)	(0.312)	(0.326)	(0.386)	(0.380)	(0.437)	(0.423)
Constant	0.103	0.068	0.160	0.047	0.182	0.130	0.287	0.086	–0.122	–0.159	–0.090	–0.268
	(0.239)	(0.232)	(0.237)	(0.259)	(0.368)	(0.365)	(0.370)	(0.382)	(0.561)	(0.534)	(0.577)	(0.544)
Sample size	254	254	232	199	254	254	228	199	254	254	222	199
R^2	0.025	0.029	0.028	0.037	0.093	0.092	0.047	0.066	0.037	0.040	0.053	0.083

Notes: (i) ***: $p \leq 0.01$; **: $0.01 < p \leq 0.05$; *: $0.05 < p \leq 0.10$. (ii) The standard errors between parentheses are based on clustering among cities in the same prefectures. (iii) "Baseline" refers to the treatment variable that takes unity when a city has not received a LAT grant in every FY between t_B and t_A. (iv) "Longer horizon" refers to the treatment variable that takes unity when a city has received a LAT grant in every FY between t_B – 3 and t_A. (v) "Movers" refer to cities whose LAT status was not constant either between t_B and t_A or between t_B – 3 and t_A.