Workload and Bureaucratic Disentitlement: Evidence from Public Assistance in Japan*

Masayoshi Hayashi

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

Abstract: In this study, we examine whether changes in workload affect the rationing behavior on welfare provision by Japanese local governments. Exploiting exogenous variation in welfare caseloads due to a large wave of municipal mergers in Japan, we investigate the impacts of workload size on rationing behavior using city-level data on welfare program applications, application withdrawals, and application rejections. Our results show that, after controlling for the existing size of caseloads, an increase in caseworker size (i.e., a decrease in workload) leads to an increase in applications and withdrawals but has no effect on rejections. Since the increase in applications exceeds the increase in withdrawals, we find that a decrease in workload leads to a higher number of accepted applications. Furthermore, heavier existing workloads are found to amplify the positive effects of caseworker size on applications and withdrawals. Our results also suggest that Japanese welfare offices may prefer informal rejections. These results lend support to the type I error explanation of bureaucratic disentitlement, as suggested in the standard literature, rather than the type II error explanation of the 'cursory assessment' hypothesis once claimed by the Japanese government.

Keywords: social assistance, workload, caseloads, caseworkers, bureaucratic disentitlement

JEL Codes: H73, H75, H77

* This study was financially supported by Grant-in-Aid for Scientific Research (C) 21K01521, (B) 20330064, and (A) 15070000033. The author would like to thank Mototsugu Fukushige, Haruko Noguchi, Yasuyuki Sato and participants at various academic meetings for their valuable comments and feedback.

1. Introduction

The heavy workload of welfare caseworkers can adversely affect the implementation of welfare programs. A smaller number of caseworkers relative to the caseload may lead to implicit rationing of their assistance. For instance, such conditions might prevent caseworkers from going beyond their job descriptions (Ridzi and London 2006) and make them more vulnerable to burnout, reducing their working efficacy (Lloyd et al. 2002). As a study of 55 caseworkers across seven welfare offices in Japan found that 28 workers (51 percent) experienced 'high burnout' (Takeda et al. 2002). Heavy workload can lead to higher staff turnover, exacerbating the workload for those who remain (Smith 2005). Caseworkers at understaffed offices may exclude qualified applications or deny existing entitlements, engaging in what is referred as 'bureaucratic disentitlement' (Lipsky 1984; Brodkin 1997). This represents a 'type I error' (rejection of the correct), where those entitled to assistance are excluded. In U.S. cities, caseworkers often discourage potential beneficiaries from applying for welfare benefits to which they are entitled (Moffitt 2003) and may even apply faulty procedures to prevent clients from receiving benefits. For example, almost half of the sanctions imposed by TANF caseworkers in the US were reversed upon client appeal (Lens 2006). This line of argument suggests that an increase in the number of caseworkers would alleviate their workloads and reduce welfare rationing.

However, an opposing claim was made in a Japanese policy debate in 2005, when the central government negotiated a cost-sharing scheme with local governments for Japan's Public Assistance (PA) system.[1] The central government sets rules for PA programs, which are then implemented by local governments. As PA benefits are provided through laborious means-testing, the central government argued that caseworkers in understaffed welfare offices were conducting 'cursory assessment' of PA applications, which led to an unnecessary increase in caseloads (Kimura 2006). In contrast to the type I error of bureaucratic disentitlement, this is a 'type II error' (acceptance of the wrong), where individuals who are not entitled are qualified. The 'cursory assessment' hypothesis then suggests that increasing the number of caseworkers would reduce workloads, lessen the occurrence of cursory assessments, and ultimately bring caseload to an appropriate level.

A larger number of caseworkers may result in fewer acceptances in a type II error environment (cursory assessment) or more acceptances in a type I error environment (bureaucratic disentitlement). In either case, it is important to examine whether understaffed welfare offices threaten the fair implementation of social programs. Despite its significance, only

[1] Public Assistance here refers to social assistance called *Seikatsu Hogo* in Japanese, which literally means "the protection (*hogo*) of daily life (*seikatsu*)."

a few empirical studies have seriously examined the impact of workload on the implementation of social programs.[2] Hainmueller et al. (2016), exploiting a large-scale pilot by Germany's employment agency, found that offices with reduced workloads increased monitoring, imposed more sanctions, and intensified search efforts while also registered additional vacancies. Meanwhile, Schmieder and Trenklee (2020), using data from the Integrated Employment Biographies of the German Social Security system, observed that caseworker teams handling larger caseloads spent less time and resources on individual workers. For Japan, Suzuki and Zhou (2007) came close to addressing the issue by regressing caseload size on the number of caseloads per caseworker. However, this analysis is less direct in evaluating the bureaucratic disentitlement or cursory assessment hypotheses, as its outcome variable is the stock value of caseloads rather than the acceptance or rejection of new applications.[3] Additionally, less attention has been given to the issue of endogeneity. If local governments allocate more human resources (caseworkers) for their programs in response to an increase in workload (caseloads), the number of caseworkers becomes endogenous. Therefore, it is crucial to account for such endogeneity.

In this study, we examine the effect of workload on welfare rationing using Japanese municipal data. There are three key advantages to using Japanese data. First, we can exploit exogenous variation in PA caseloads resulting from municipal boundary reforms in the mid-2000s. Japan's system of local government is two-tiered, with municipalities (cities, town, villages) forming the first tier and prefectures forming the second. National law requires cities and prefectures to establish welfare offices to implement PA programs, while towns and villages (TVs) are not required to do so, as PA programs for residents in TV areas are managed by prefectural welfare offices (with a small number of exceptions). Therefore, when a city merges with TVs, it begins to cover PA recipients in the ex-TV areas who were previously served by the prefectural offices. As a result, cities merging with TVs experience exogenous increases in the number of PA recipients. This exogenous variation should provide a valid instrument for estimating the effect of increasing PA caseworkers in our analysis.

---

[2] Numerous studies have explored the factors influencing welfare caseloads, though most focusl on aspects other than caseworkers' workload. The majority of them examine the United States, with similar analyses conducted in Canada (Spindler and Gilbreath 1979), Sweden (Gustafsson 1984), Spain (Ayala and Pérez 2005), and Japan (Suzuki and Zhou 2007). Except for Brehm and Saving (1964), US studies were largely motivated by the sharp increase in caseloads observed in the early 1990s, followed by an abrupt decline after 1994, which coincided with a series of welfare reforms at both the state and federal levels. As a result, the US literature primarily investigates the effects of economic factors—such as income levels and unemployment—along with the impact of changes in welfare programs (Schiller and Brasher 1993; Johnson et al. 1994; Schiller 1999; Ziliak et al. 2000; Blank 2001; Huang et al. 2004; Moffitt 2003; Cadena et al. 2006; Danielson and Klerman 2008). Additionally, research explores other influential factors, including at-risk populations (Conte et al. 1998), sluggish adjustments in welfare participation (Figlio and Ziliak 1999; Ziliak et al. 2000), and regional labor market conditions (Lee et al. 2002; Lewis and Henry 2004; Hill and Murray 2008)..

[3] This study argues that more caseworkers imply more resources to encourage clients to move away from reliance on welfare benefits and toward self-sufficiency, thereby reducing the number of welfare recipients.

Second, the institutional aspect of the Japanese programs provides another advantage. Local governments in Japan implement PA programs according to the uniform rules set by the central government. They cannot alter the level of assistance or other policy parameters in the system. Therefore, the endogeneity of such policy parameters (Mayer 2000) is less of a concern in our estimation. Additionally, this uniformity allows us to utilize data of all cities within the country without worrying about issues arising from differences in the assistance systems among subnational regions. This contrasts with US studies that use data at the implementers' level (counties). To avoid issues stemming from interstate differences, these studies tend to use small samples, consisting of units in a single state (Grubb 1984, Lee et al. 2002, Kerman and Haider 2004, Hill and Murray 2008).

Third, by using unpublished administrative data at the welfare office level from the *Report on Social Welfare Administration and Services* compiled by the Japanese Ministry of Labour, Health and Welfare (MLHW), we use more detailed output variables that describe welfare rationing behavior, rather than relying solely on the data for PA caseload size. We examine three variables related to assistance rationing at the city level, which are expected to be influenced by workload: (1) the number of applications for PA programs, (2) the number of application withdrawals, and (3) the number of application rejections. We include 'applications' because welfare offices often conduct intake interviews with potential applicants before the formal submission of their applications. Anecdotal evidence, including reports by newspapers and journalists, suggests that welfare offices often use these intake interviews to expel potential recipients who are on the margins of eligibility but are still entitled to PA benefits. Similarly, it could also be argued that, even among those who are 'eligible' to apply, welfare offices may try to persuade some of them to withdraw their applications. This type of rationing behavior may or may not be facilitated by the availability of resources at welfare offices. The same logic should also be applied to the rejection of applications.

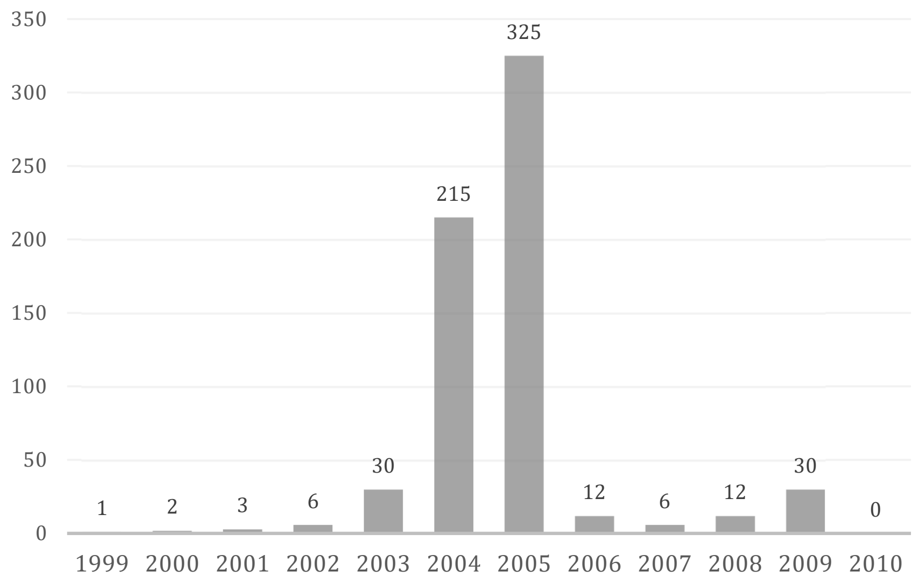
The remainder of this paper is organized as follows. After a brief description of the Japanese institutional background, Section 2 explores changes in PA caseloads in cities that have merged with TVs and the subsequent changes in the number of caseworkers. Section 3 then sets up regression models to explore the effect of caseworker size on the rationing behavior of welfare offices. After discussing the estimation results, Section 4 concludes.

2. The effect of PA caseload on caseworker size

2.1. Municipal mergers and changes in PA caseloads

To obtain exogenous variation in PA caseloads, we exploit the wave of municipal mergers that took place in Japan in the mid-2000s. The wave was triggered by a policy shift in 1999, when legislation was enacted to promote fiscal decentralization which emphasizes the role of municipalities in providing public services. Recognizing that many municipalities were too small to manage decentralized functions effectively, the central government incentivized mergers through generous fiscal and administrative support. As a result, a significant number of mergers occurred, reducing the number of municipalities by 47% from 3,229 at the end of FY1999 to 1,727 at the end of FY2010. As illustrated in **Figure 1**, these mergers were most concentrated in FY2004 and FY2005, peaking in FY2005. After that, incentives for mergers were significantly reduced, and the promotion campaign officially ended at the end of FY2009.

Figure 1. Number of municipal mergers: 2000–2010

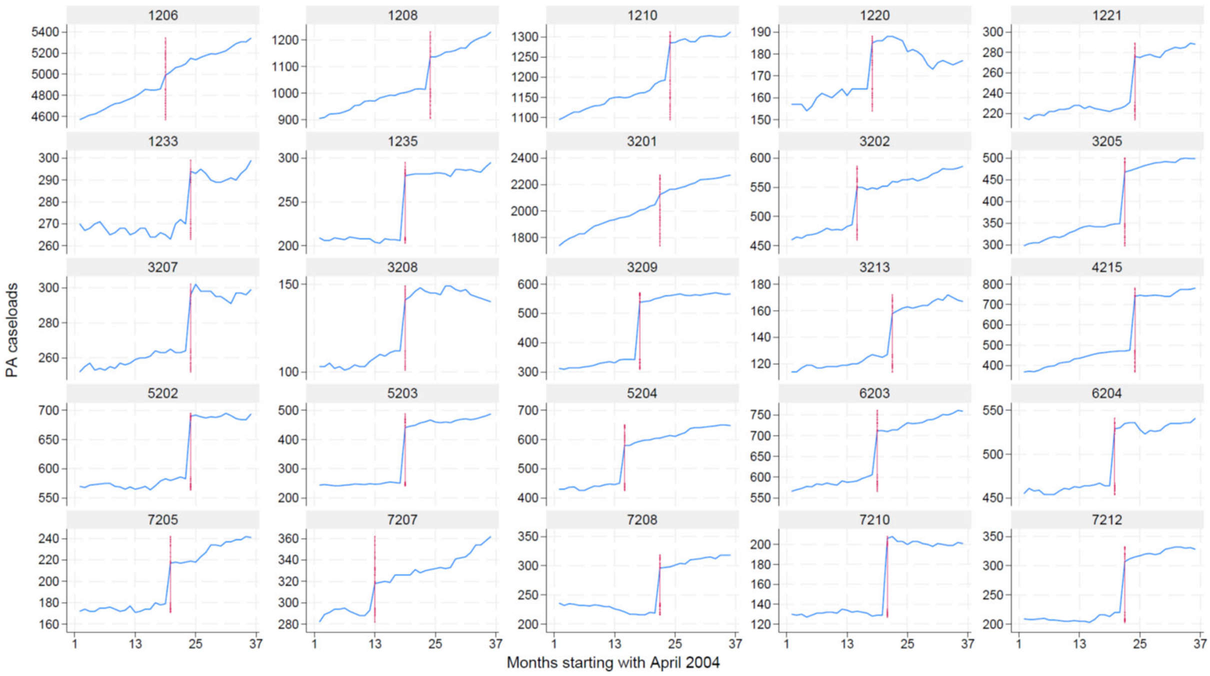


Source: Ministry of Internal Affairs and Communications.

The key source of exogenous variation is cities that merged with TVs. As explained in the Introduction, such mergers increase the number of PA recipients in the merged city, as the original city's PA program expands to include recipients in former TV areas who were previously covered by the prefectural program. Since municipal decisions to merge are orthogonal to the issues handled by welfare offices, these changes in the number of caseloads can be considered exogenous. At the same time, we expect an increase in the number of PA caseworkers in the original city, as national law recommends that local governments maintain a ratio of one caseworker per 80 recipients. However, since this is merely a recommendation rather than a strict requirement, we anticipate variation in municipal responses depending on their specific characteristics.

Among the 786 cities that existed on April 1, 2010, the reform enabled 341 cities to merge with TVs during the 2000s. We focus on mergers in FY2005, the fiscal year with the highest number of municipal mergers during this period.[4] Excluding cities with obvious data anomalies, we obtain a sample of 125 cities that merged with TVs in FY2005 (April 2005 to March 2006). The monthly PA caseloads for each city are displayed in panels in **Figure 2**, covering the period from April 2004 to March 2007. In each panel, the vertical line marks the month of the merger, while the horizontal axis represents the number of months since April 2004. The panels clearly illustrate abrupt changes in monthly PA caseloads in the months when mergers took place.

Figure 2. Monthly caseload trends for cities merged with TVs in FY2005



[4] In our preliminary analysis, we conducted an analogous estimation using FY2004 mergers with FY2004 and FY2005 data. For this exercise, we used a sample of 452 cities, comprising 82 cities that merged with TVs in FY2004 (April 2004 to March 2005) and the unmerged cities, identical to those used in this study. However, the instruments in this analysis were found to be quite weak, possibly be due to the small share of merged cities (0.181=81/450 for FY2004, compared to 0.254=126/496 for FY2005). As a result, the IV estimates using FY2004/FY2005 data are all statistically insignificant.

Figure 2. Monthly caseload trends for cities merged with TVs in FY2005 (Continuing)

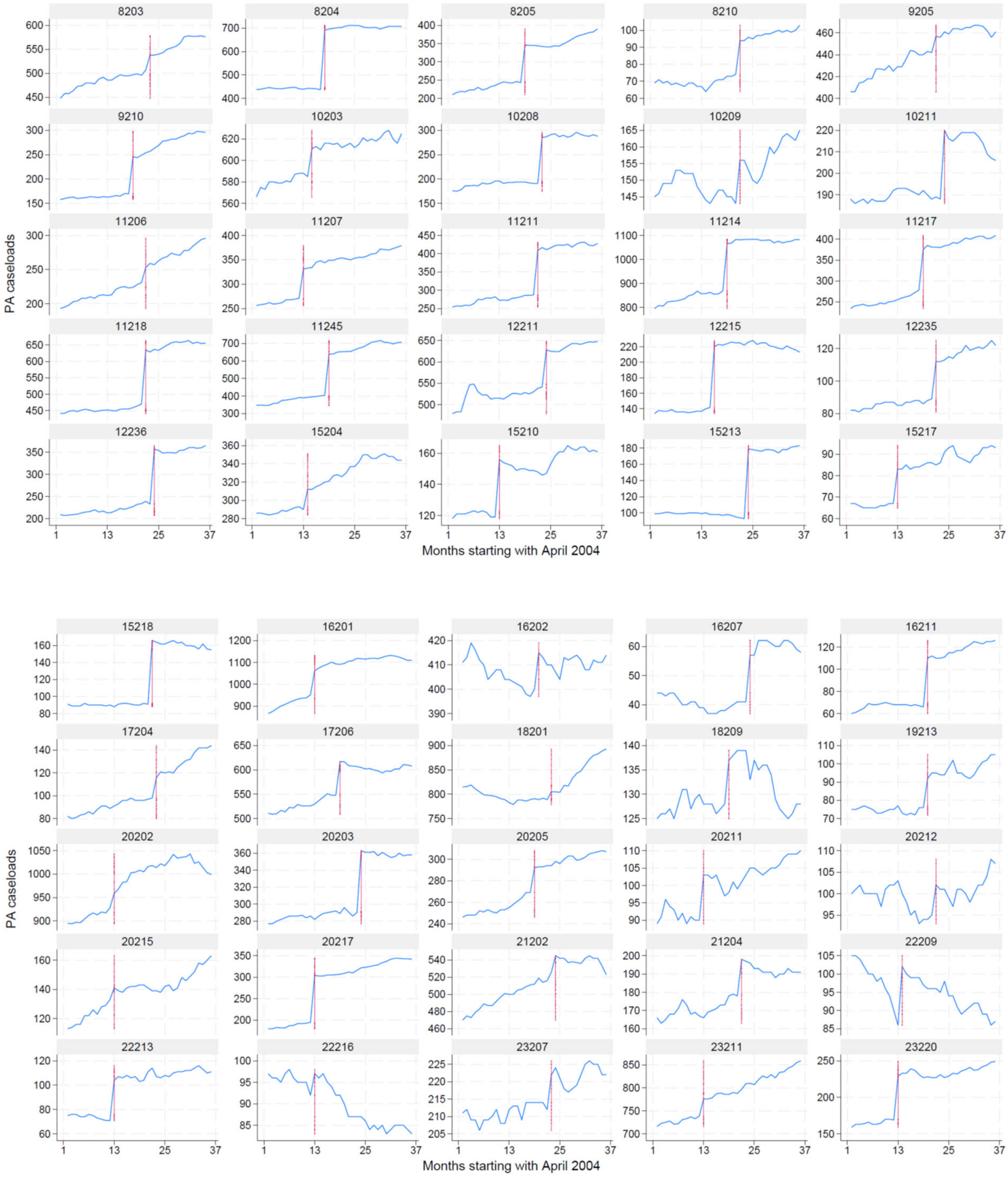
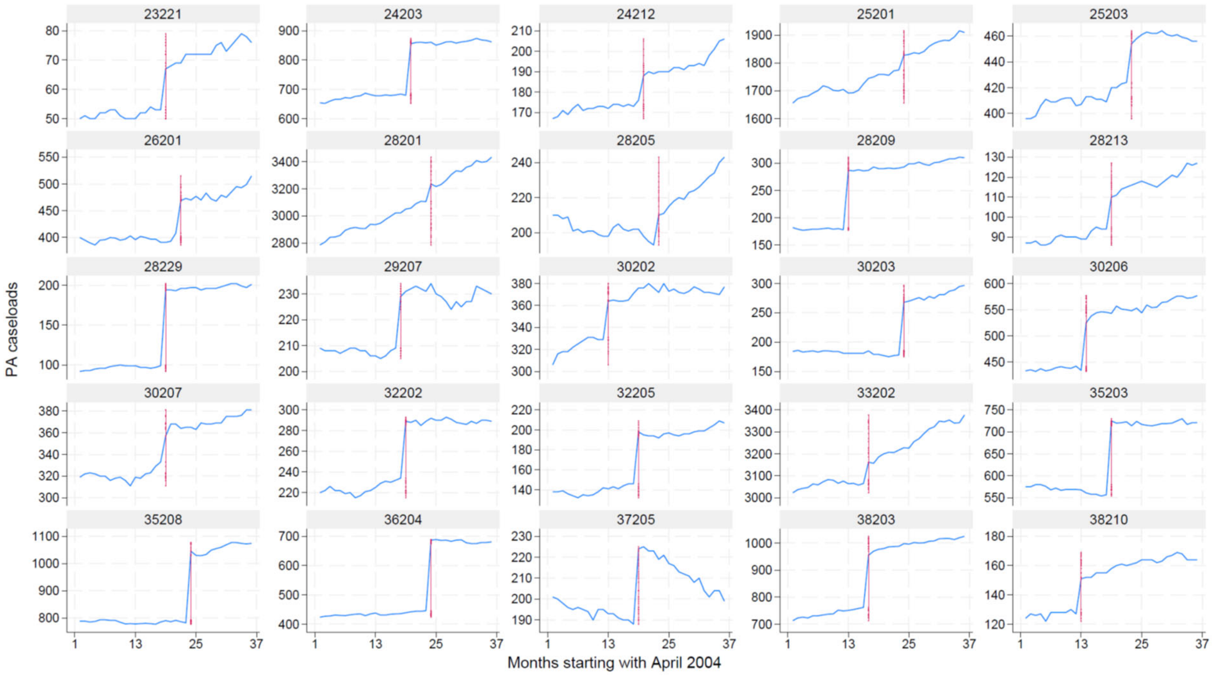
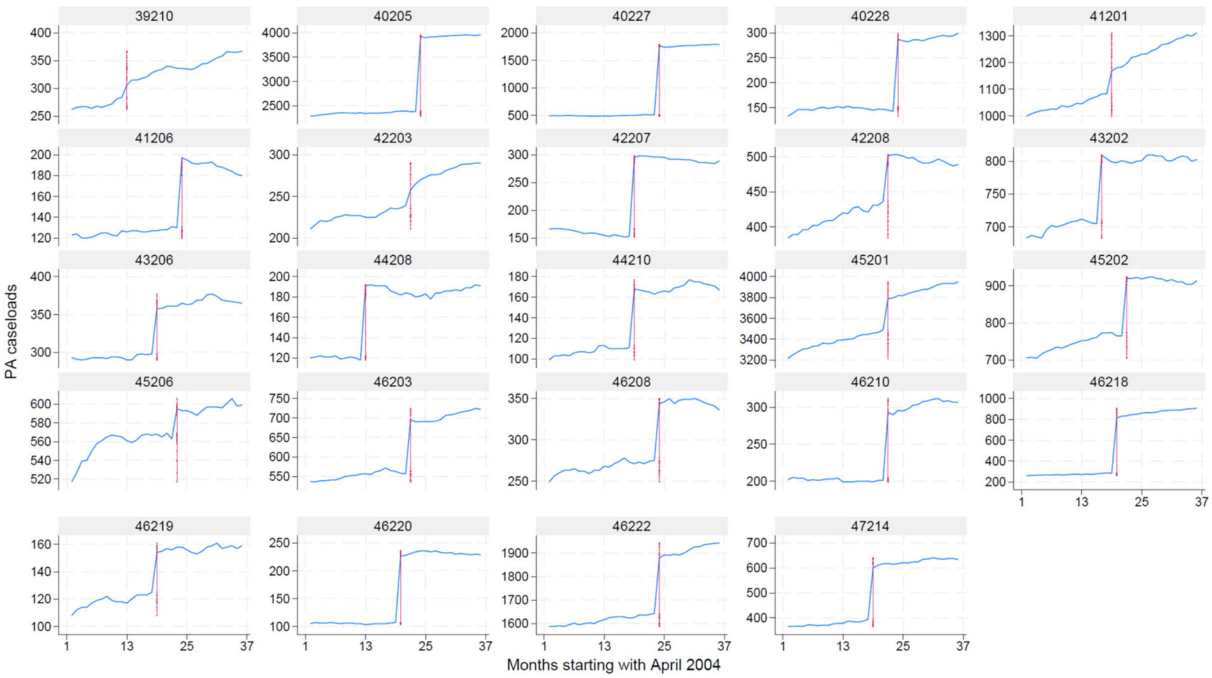


Figure 2. Monthly caseload trends for cities merged with TVs in FY2005 (Continued)



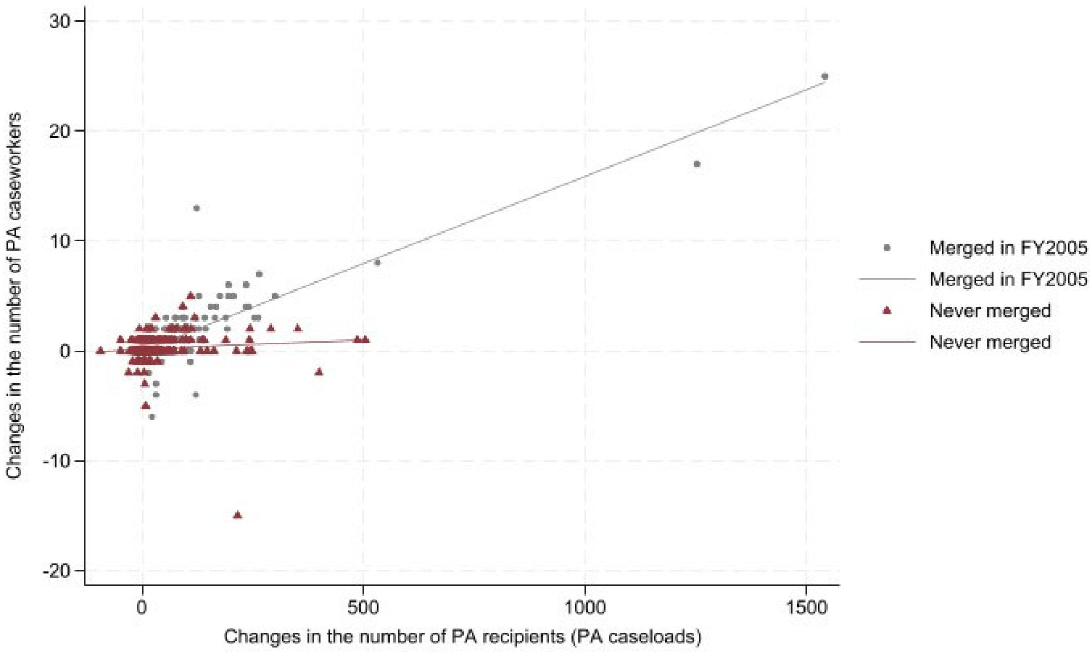


Source: Ministry of Health, Labour and Welfare. Report on Social Welfare Administration and Services.

2.2. Increased PA caseloads and resulting changes in PA caseworkers

Our sample consists of these 124 merged cities in **Figure 2** and the other 370 cities that did not merge in the 2000s. **Figure 3** plots and fits the changes in PA caseloads (horizontal axis) and those in PA caseworkers (vertical axis). The dots and their fitted line represent the changes in those cities that merged with TVs in FY2005. For these merged cities, the changes in PA caseloads are calculated as the differences between the caseloads in the month of the merger and those in the month immediately preceding it. Since the number of PA caseworkers is measured at the start of the fiscal year (i.e., on April 1), differences are taken between the first days of FY2005 and FY2006. Despite some erratic observations with negative caseworker changes, the dots indicate a positive correlation between the two variables. Meanwhile, the triangles and their fitted line represent cities that never merged during the 2000s. Changes in their caseloads are measured as differences between the annual average values for FY2005 and FY2006. These triangles suggest that the correlation appears weaker and flatter among unmerged cities than among merged cities. Overall, the figure suggests that changes in PA caseloads due to mergers with TVs could serve as a valid instrument for our estimation.

Figure 3. Correlation between changes in caseloads and caseworkers



Notes: The dots and their fitted line represent the changes in cities that merged with TVs in FY2005, while the triangles and their fitted line represent cities that had never merged during the 2000s. The changes in PA caseloads for the merged cities are calculated as the differences between the caseloads in the month of the merger and those in the month immediately preceding the merger. The number of PA caseworkers is at the start of the fiscal year, and the differences are calculated between the first days of FY2005 and FY2006.

Using this group of cities as the sample, we regress changes in the size of caseworkers on changes in PA caseloads to more closely observe their association. Our regression model begins with a linear model with undifferenced variables and unobserved heterogeneity c_i^x:

$$x_{it} = \alpha \cdot z_{it-1} + \sum_k \gamma_k^x \cdot w_{k.it} + c_i^x + \mu_t^x + u_{it}^x$$

for $t =$ FY2005 and FY2006. When differenced, the regression model above yields:

$$\Delta x_{it} = \alpha \cdot \Delta z_{it-1} + \sum_k \gamma_k^x \cdot \Delta w_{k.it} + \delta^x + \epsilon_{it}^x \tag{1}$$

for $t =$ FY2006. Note that we cannot use a panel of differenced data because we can only validly investigate the effect of the caseworker change in a single fiscal year (FY2006), which immediately followed the fiscal year of the exogenous change in PA caseloads (FY2005).

Also note the following. First, Δx_{it} represents the annual changes in the number of caseworkers from the first day of FY2005 to that of FY2006. Second, Δz_{it-1} measures the change in caseload caused by municipal mergers in FY2005. The value of Δz_{it-1} for a merged city is the difference in caseloads around the vertical line in each panel of **Figure 2**, with its size representing the change from the month of the merger to the month immediately following it. For unmerged cities, the value of this variable is zero. Third, $w_{k.it}$'s are covariates consisting of the size of caseloads in the last month of the previous fiscal year, the share of female population, the share of elderly (65+) population, population size, the number of households, and fiscal capacity index. The fiscal capacity index measures the ratio of a local government's revenue capacity to its spending needs. These covariates also serve as controls for the IV estimation in the subsequent section. Fourth, $\delta^x \equiv \Delta\mu_t^x$ represents the differenced year effects, which are absorbed into the constant term as we use a cross-section of the differenced data for $t =$ FY2006. Lastly, $\epsilon^x \equiv \Delta u_{it}^x$ is error term. The sample statistics and data sources are listed in the Appendix.

As we argue that the caseload changes caused by mergers (Δz_{it-1}) are orthogonal to the error term, the OLS estimation of α provides a valid estimate. We first estimate a version of Eq. 1 that excludes all covariates ($\gamma_k^x = 0$ for all k). As shown in the second column of **Table 2**, the coefficient is estimated at 0.0156 and is statistically significant. This indicates that, on average, one additional caseworker is assigned when the PA caseload increases by 64 (= 1/0.0156). We then estimate Eq. 1 including all covariates, which are presented in the last column of the table. Adding covariates reduces the estimated value to 0.0140 which, while statistically significant, implies that one caseworker is added when the caseload increases by 71 (= 1/0.0139).

The Japanese government recommends, rather than requires, that local governments set the 'standard' caseload-to-caseworker ratio at 80 PA recipient households per caseworker. The results above show that, on average, cities that merged with TVs in FY2005 increased their number of PA caseworkers beyond the level recommended by this standard. However, this is expected, as the denominator of the standard ratio includes not only PA caseworkers but also other caseworkers assigned to different programs implemented by welfare offices.

Table 1. Effect on PA caseworkers

	Without covariates	With covariates
Changes in PA caseloads due to mergers	0.0156*** (0.001)	0.0139*** (0.002)
F-value	405.0	63.9
R^2	0.588	0.599
Sample size	494	494

Notes: (i) ***: $p \leq .01$; **: $.01 < p \leq .05$; *: $.05 < p. \leq 10$. (ii) Standard errors are in parentheses.

3. Effect of workload on assistance rationing

3.1. Outcome variables for assistance rationing

We use three variables related to assistance rationing at the city level, which are explained by caseworker size: the numbers of (1) the number of applications to PA programs, (2) the number of application withdrawals, and (3) the number of application rejections. As we discussed in the Introduction, we consider 'applications' because welfare offices can use intake interviews to expel potential recipients before they formally submit their applications. Additionally, even among those who have submitted applications, welfare offices might also try to persuade some applicants to withdraw.

3.2. Regression model

To estimate the effects of caseworker size, we begin with the following regression model, using a generic expression y for each of the three variables:

$$y_{it} = \beta \cdot x_{it-1} + \sum_k \gamma_k^y \cdot w_{k.it} + c_i^y + \mu_t^y + u_{it}^y,$$

for $t =$ FY2005 and FY2006. Notice that the explanatory variables are already defined, and the parameters and errors are similarly defined as in Eq. 1. Following the same approach as in Eq. 1, we take difference of the model to obtain the following:

$$\Delta y_{it} = \beta \cdot \Delta x_{it} + \sum_k \gamma_k^y \cdot \Delta w_{k.it} + \delta^y + \epsilon_{it}^y \qquad (2)$$

for $t = \text{FY2006}$, where δ^y and ϵ_{it}^y are defined analogously to in Eq. 1. Again, we cannot utilize a panel of differenced data, as we can only examine the effect of the caseworker change in FY2006, which immediately followed FY2005 when the exogeneous change in caseloads occurred. The sample for estimating Eq. 2 is identical to that used for the estimation of Eq. 1.

Note the following. First, Δx_{it}measures changes in PA caseworkers from the beginning of FY2005 to that of FY2006, which is instrumented by $\Delta z_{i\text{FY2005}}$, which measures changes in PA caseloads caused by mergers with TVs in FY2005. Note that this variable takes a value of zero for cities that did not experience mergers during the 2000s.

Second, while $y_{i\text{FY2005}}$ in Δy_{it} for cities without mergers takes on recorded annual values, $y_{i\text{FY2005}}$ for cities with mergers considers only the relevant values in the months before their mergers in FY2005. We calculated their monthly averages before the mergers and multiplied them by 12 to obtain their annual equivalents. By doing so, we can associate $y_{i\text{FY2005}}$ with $x_{i\text{FY2005}}$, the size of PA caseworkers at the beginning of FY2005, i.e., the period before the mergers.

Third, it is necessary to include the existing caseload size as a covariate in Eq. 2 to determine the effect of workload. 'Workload' can be defined as the amount of work required for a single caseworker to complete his/her designated tasks (Strolin et al. 2007). Although it is difficult to precisely measure workload using this definition, we can reasonably approximate it by using the size of caseloads per caseworker or an 'average workload' in a locality. To account for the effect of average workload, we included the size of PA caseloads in the last month of the previous fiscal year in the set of the covariates ($w_{k,it}$) to represent the existing caseload size handled by a welfare office. By doing so, we can interpret changes in caseworker size as corresponding to changes in average workload in the opposite direction. If $\beta > 0$, it then implies that a higher (lower) workload would lead to a smaller (larger) output. Conversely, if $\beta < 0$, it implies that a higher (lower) workload would lead to a larger (smaller) output.

Fourth, including caseload size as a regressor in Eq. 2 introduces another endogeneity issue. Since the last month of fiscal year t is March of calendar year $t+1$, we obtain $\Delta w_{1,i\text{FY2006}} = z_{i\text{March 2006}} - z_{i\text{March 2005}}$. Given that $y_{i\text{FY2005}}$ plausibly affects $z_{i\text{March 2006}}$, $\Delta y_{i\text{FY2006}} = y_{i\text{FY2006}} - y_{i\text{FY2005}}$ would also influence $\Delta w_{1,i\text{FY2006}} = z_{i\text{March 2006}} - z_{i\text{March 2005}}$, leading to a reserve causality issue. While our instrument $\Delta z_{i\text{FY2005}}$ clearly affects $\Delta w_{1,i\text{FY2006}}$, we need another instrument to achieve (just) identification in IV estimates. To address this, we employ an Anderson-Hsiao-type instrument, using twice-lagged values of $\Delta w_{1,i\text{FY2006}}$: $\Delta w_{1,i\text{FY2004}} = z_{i\text{March 2004}} - z_{i\text{March 2003}}$. This variable remains uncorrelated with both $y_{i\text{FY2006}}$ and $y_{i\text{FY2005}}$ if serial correlation in y_{it} is less than second order. Note that we cannot perform Hansen's J test since our IV regression is just identified, using two instruments ($\Delta z_{i\text{FY2005}}, \Delta w_{1,i\text{FY2004}}$) for two endogenous regressors ($\Delta x_{i\mathrm{FY2006}}$, $\Delta w_{1,i\mathrm{FY2006}}$).

Lastly, Eq. 2 accounts for unobserved heterogeneity c_i^y which captures several important factors noted in the literature. For instance, caseworkers may adopt collective values shared within their organizations (Keiser and Soss 1998). Community attitudes also play a crucial role as they can discourage eligible individuals from applying for welfare or cause caseworkers to take stricter positions on eligibility assessments (Grubb 1984; Weissert 1994). Since these shared values are likely to remain constant over short periods but vary across cities, unobserved heterogeneity absorbs their effects beyond those of other locality-specific factors that are stable over time.

3.3 Effects of workload

Table 2 presents the IV estimates for the effects of PA caseworker size and the existing size of PA caseloads on the three rationing outcomes: applications, withdrawals, and rejections. For comparison, we also report the corresponding OLS estimates to the right of their IV counterparts. Since the error term is likely to be nonspherical with the cross-section of differenced data, the standard errors (in parentheses) are obtained with the heteroskedasticity-consistent covariance matrix estimator. As shown in the last four rows of the table, our instruments pass the weak instrument test proposed by Montiel Olea and Pflueger (2013) at the significance level of $\alpha = 0.05$ and a desired threshold of $\tau = 0.10$.[5] Moreover, the table reveals substantial differences between the OLS and the IV estimates, supporting the validity of using the IV estimator.

The IV estimation reveals statistically significant effects of PA caseworker size, except for rejections. The estimate for applications suggests that increasing the number of caseworkers by one, while holding the existing caseload size constant, would lead to an annual increase of 24 applications. As previously discussed, welfare offices may use intake interviews to discourage potential applicants from formally submitting their applications. An increase in caseworkers may reduce the emphasis on such implicit rationing, allowing staff to focus on other tasks.

These other activities may be reflected in the number of withdrawals and rejections. Our estimates indicate that caseworker size affects withdrawals but not rejections. Since applications and withdrawals are likely to be positively correlated, an increase in applications due to a larger size of caseworkers also results in more withdrawals. Additionally, the cursory assessment hypothesis suggests that with more caseworkers available, welfare offices might allocate more resources to scrutinizing applicants, potentially leading to more rejections. However, our findings contradict this expectation. One possible explanation is that welfare offices may prefer to avoid

[5] For this test, we used a Stata postestimation routine **weakivtest**, developed by Pflueger and Wang (2015).

formal rejections by informally persuading applicants to withdraw their applications. This pattern —an increase in applications and withdrawals, but no significant effect on rejections—suggests that Japanese welfare offices tend to rely on informal case resolution. Nonetheless, the rise in withdrawals (six additional withdrawals) was not substantial enough to offset the increase in applications (24 additional applications). Since an increase in caseworker size does not lead to more rejections, it ultimately contributes to more acceptances of PA recipients.

The IV estimates for the existing size of PA caseloads align with the effects of caseworker size. Holding caseworker size constant, an increase in existing caseloads implies a higher workload. The negative coefficients for applications and withdrawals correspond to the positive coefficients for caseworker size in both cases. Additionally, neither the existing size of caseloads nor the number of caseworkers has a significant effect on rejections.

Table 2. Effects on rationing behavior

		Applications		Withdrawals		Rejections	
		IV	OLS	IV	OLS	IV	OLS
Caseworker size		24.37^{***} (7.56)	4.71^{*} (2.56)	5.72^{***} (1.51)	1.39^{*} (0.83)	0.05 (0.69)	0.30^{**} (0.16)
Existing caseload size		-0.25^{**} (0.11)	0.01 (0.05)	-0.06^{***} (0.02)	-0.004 (0.02)	0.005 (0.01)	0.005 (0.003)
Effective F ($\alpha = 0.05$)		19.44		19.44		19.44	
Critical values for $\alpha = 0.05$	$\tau = 0.10$	18.05		18.01		17.99	
	$\tau = 0.20$	11.97		11.94		11.93	
	$\tau = 0.30$	9.68		9.65		9.64	
Sample size		494	494	494	494	494	494
#cities merged with TVs		124	124	124	124	124	124

Notes: (i) ***: $p \leq .01$; **: $.01 < p \leq .05$; *: $.05 < p. \leq 10$. (ii) Standard errors are in parentheses. (iii) The sample size is 496, including 126 cities that merged with TVs. (iv) The last four rows present statistics from the weak instrument test by Montiel Olea and Pflueger (2013).

3.4 Effects of workload with different intensity

We expect the negative effect of an increasing number of caseworkers on rationing behavior to become more pronounced as the workload of the existing caseworkers intensifies. To explore these differences, we split the sample based on a threshold for work intensity. As noted in Section 2, the Japanese government sets a 'standard' workload for welfare offices in cities at 80 recipient households per caseworker (RPC). Therefore, we divide the sample into two subsamples: one with $RPC > 80$ and the other $RPC \leq 80$.

Table 3 presents the results for each of the two subsamples. When estimated using the subsample with $RPC > 80$, the instruments passed the Montiel Olea-Pflueger test for weak instruments at a significance level of $\alpha = 0.05$, although with a relatively high desired threshold of $\tau = 0.20$. In contrast, when estimated using the subsample with $RPC \leq 80$, we can lower the threshold level to $\tau = 0.10$ and still pass the test. Analogous to **Figure 3**, we also plot and fit the changes in caseloads and caseworkers for the two subsamples in two panels of **Figure 4**. A visual inspection of the dots and triangles suggests that the correlation appears weaker among unmerged cities than among merged cities, further validating our use of changes in PA caseloads due to mergers with TVs as one of the instruments.

The coefficients for $RPC > 80$ are nearly identical to those from the full sample (IV estimates in **Table 2**), although current caseloads have a statistically insignificant effect on applications and a less significant effect on withdrawals. Meanwhile, neither caseworker size nor existing caseload size for $RPC \leq 80$ significantly affect the applications, which is to be expected, as we are examining welfare offices with relatively low workloads. In contrast, the effect on withdrawals for $RPC \leq 80$ is statistically significant. It is difficult to interpret the lack of significance for applications in this case. While offering a definitive interpretation is challenging, low workloads may allow caseworkers to spend more time consulting potential recipients and making more of an effort to dissuade them from proceeding with the formal application process.

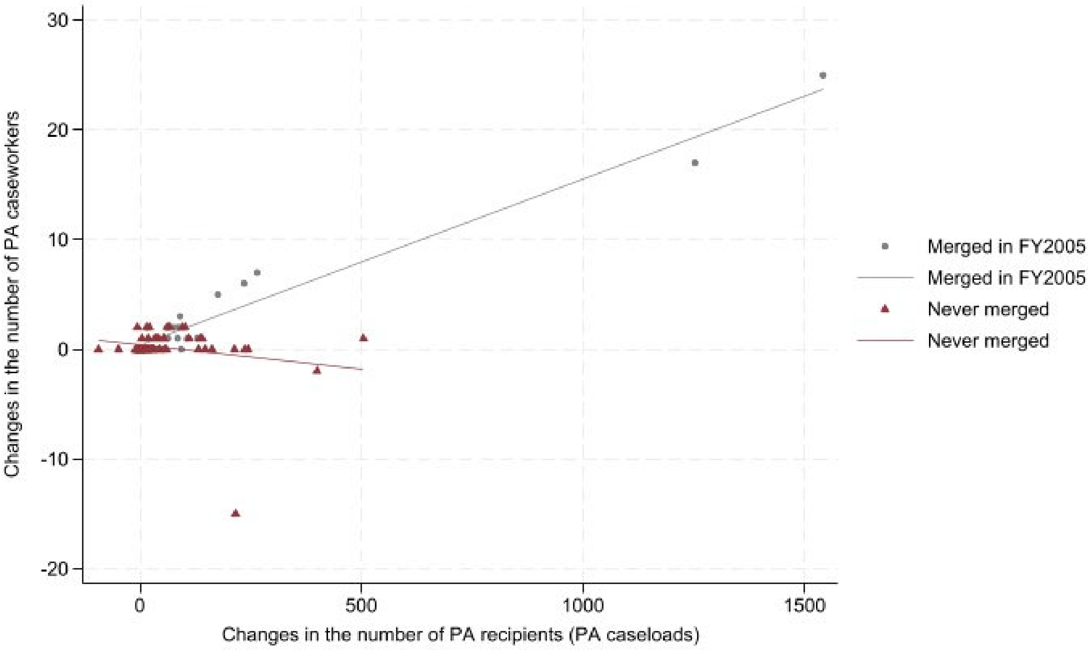
Table 3. Effect on rationing behavior with different workload

		Applications		Withdrawals		Rejections	
PA recipient households per caseworker (RPC)		> 80	≤ 80	> 80	≤ 80	> 80	≤ 80
Caseworker size		24.65** (11.51)	19.86 (13.39)	6.03** (2.66)	4.82*** (1.52)	0.23 (0.65)	0.27 (1.53)
Existing caseload size		−0.26 (0.17)	−0.20 (0.20)	−0.06* (0.04)	−0.05*** (0.02)	0.005 (0.01)	0.01 (0.02)
Effective F ($\alpha = 0.05$)		13.07	14.32	13.07	14.32	13.07	14.31
Critical values for $\alpha = 0.05$	$\tau = 0.10$	15.99	10.83	15.85	9.56	15.77	9.73
	$\tau = 0.20$	10.70	7.54	10.62	6.80	10.57	6.90
	$\tau = 0.30$	8.70	6.29	8.64	5.74	8.60	5.80
Sample size		87	407	87	407	87	407
#cities merged with TVs		24	100	24	100	24	100

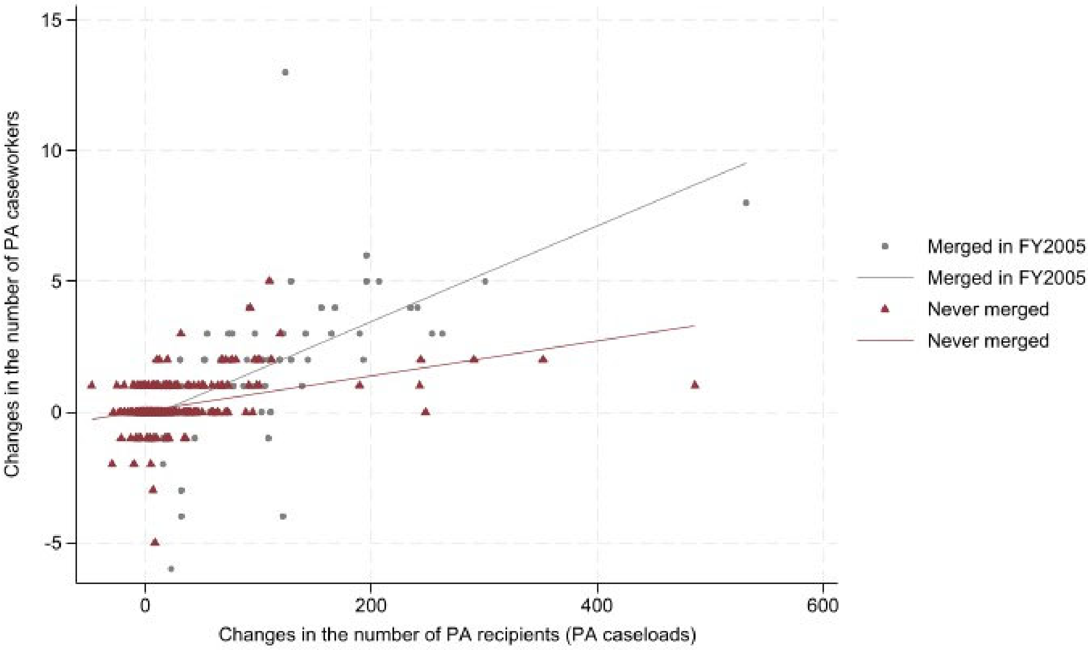
Notes: (i) ***: $p \leq .01$; **: $.01 < p \leq .05$; *: $.05 < p. \leq 10$. (ii) Standard errors are in parentheses. (iii) The sample size for $RPC > 80$ is 89, including 25 cities that merged with TVs. (iv) The sample size for $RPC < 80$ is 407, including 101 cities that merged with TVs. (v) The last four rows show the statistics for the test for weak instruments by Montiel Olea and Pflueger (2013).

Figure 4. Correlation between changes in caseloads and caseworkers by workload intensity

(a) High workload tensity: PA recipients per caseworker > 80



(b) Non-high workload Intensity: PA recipients per caseworker ≤ 80



Notes: The dots and their fitted line represent the changes in cities that merged with TVs in FY2005, while the triangles and their fitted line represent cities that had never merged during the 2000s. The changes in PA caseloads for the merged cities are calculated as the differences between the caseloads in the month of the merger and those in the month immediately preceding the merger. The number of PA caseworkers is at the start of the fiscal year, and the differences are calculated between the first days of FY2005 and FY2006.

The coefficients on caseworker size and existing caseload size for $RPC \leq 80$ are smaller than those obtained for $RPC > 80$. In other words, an increase in the number of caseworkers has a greater impact on applications and withdrawals when the existing workload is higher. Meanwhile, as with the results with the full sample in **Table 2**, the effects on rejections remain negligible and statistically insignificant both for $RPC > 80$ and $RPC \leq 80$. This again suggests that when a welfare office identifies flaws in applications, it may prefer to avoid officially recording rejections and, instead, informally persuade them to withdraw, regardless of the workload intensity.

4. Concluding remarks

In this study, we examined whether changes in workload affect rationing behavior in Japanese welfare offices. We exploited data from a large wave of municipal mergers in the 2000s, during which cities that merged with towns and villages experienced exogenous increases in their caseload sizes. We used three indicators: applications for PA programs, application withdrawals, and application rejections. By obtaining instruments from municipal mergers in FY2005, we estimated the impact of caseworker size on these three outputs with FY2005/FY2006 data. In doing so, we controlled for the existing size of caseloads in welfare offices to interpret an increase in caseworker size as a reduction in workload.

Our results showed that an increase in caseworker size (i.e., a decrease in workload) led to an increase in PA applications and withdrawals but had no effect on rejections. Since the increase in applications exceeded the increase in withdrawals, we concluded that, holding existing caseloads constant, a larger caseworker size led to a higher number of accepted applications. Furthermore, we found that heavier existing workloads tend to amplify the positive effects of caseworker size on applications and withdrawals.

These findings lend stronger support to the type I error explanation of bureaucratic disentitlement, as suggested in the standard literature, rather than the type II error explanation of the 'cursory assessment' hypothesis once claimed by the Japanese government. Additionally, our results suggest that Japanese welfare offices may prefer *informal* rejections, as indicated by the observed effects on applications and withdrawals but not on formal rejections.

Of course, this study has limitations. In particular, our analysis is based on data from FY2005 and FY2006, meaning that our conclusions may not be externally valid. As is often the case with natural experiments that rely on historical events, it is uncertain whether our findings would hold in settings that differ significantly from the FY2005-FY2006 context examined in this study.

Appendix: Data description

Table A1 presents the summary statistics for the variables used in this study. While our estimation is based on differenced data, the values listed in the table are in levels, i.e., before differencing. We provide the following details. First, we obtain monthly data for PA caseloads (PA recipient households) and the numbers of PA applications, withdrawals, and rejections from the *Report on Social Welfare Administration and Services* compiled by the Ministry of Health, Labour and Welfare. Confidential administrative data are recorded at the welfare office level. For cities with multiple welfare offices, we aggregate office-based data into city-level data. Additionally, when we estimate regression models, monthly data on applications, withdrawals, and rejections are aggregated into annual totals.

Table A1. Sample statistics

Nobs=496	FY2005				Fiscal Year 2006			
	Mean	St dev.	Min.	Max.	Mean	St dev.	Min.	Max.
Applications i	132.20	198.09	2	2,005	129.51	187.79	2	1,678
Withdrawals I	7.77	16.62	0	199	7.53	13.52	0	147
Rejections I	4.55	8.48	0	101	5.85	9.18	0	74
PA caseworkers ii	7.74	10.15	0	82	8.30	10.57	0	84
Caseworkers (PA + others) ii	12.63	15.09	0	113	13.46	16.08	0	127
Caseload changes by mergers i	25.74	102.19	0	1,542	0.00	0.00	0	0
Mergers in FY2005 (binary) iii	0.25	0.44	0	1	0.25	0.44	0	1
No mergers (binary) iii	0.75	0.44	0	1	0.75	0.44	0	1
PA caseloads i	676	1,055	19	9,484	730	1,118	24	10,011
Ratio of female population (%) iv	51	1	48	55	0.51	1	47	55
Ratio of elderly population (%) iv	21	5	10	40	0.22	5	10	52
Population (in thousands) iv	107.30	100.88	5.32	662.60	107.32	101.23	5.12	662.57
Households (in thousands) iv	42.19	42.09	2.75	280.64	42.70	42.67	2.65	283.31
Fiscal capacity index (%) v	70	27	11	172	71	28	12	172

Sources: (i) Report on Social Welfare Administration and Services (Fukushi Gyosei Hokokurei) by the Ministry of Health, Labour and Welfare. (ii) Annual Survey of Municipal Human Resources (Chiho Kokyo Dantai Teiin Kannri Chosa) by the Ministry of Internal Affairs and Communication. (iii) List of Municipal Mergers Since FY1999 (Heisei Juichi Nendo Iko no Shi-cho-son Gappei no Jisseki) by the Ministry of Internal Affairs and Communication. (iv) System of Social and Demographic Statistics (SSDS: Shyakai Jinko Tokei Takei) by the Statistics Bureau. (v) Annual Survey of Municipal Finance (Shi-cho-son Kessan Jokyo Shirabe) by the Ministry of Internal Affairs and Communication.

Second, the number of caseworkers was measured on the first day of each fiscal year (April 1) in each municipality. We use stock data at the beginning of the fiscal year rather than the annual average of daily counts, as the latter is unavailable. However, a Japanese study (Nakajima and Arakawa 2004) suggests that municipalities typically set their caseworker sizes at the start of a fiscal year and maintain them throughout the period.

Third, we include as many relevant covariates as possible. The data availability at the city level is limited, except for census years (in this case, only FY2005). As a result, we use a limited set of covariates, comprising population, the number of households, the ratios of women and elderly (+65) to the total city population, and fiscal capacity index.

References

Ayala, L., Pérez, C., 2005. Macroeconomic conditions, institutional factors and demographic structure: What causes welfare caseloads? Journal of Population Economics 18 (3), 563–581.

Blank, R.M., 2001. What causes public assistance caseloads to grow? Journal of Human Resources 36 (1), 85–118.

Brehm, C.T., Saving, T.R., 1964. The demand for general assistance payments. American Economic Review 54 (6), 1002–1018.

Brodkin, E.L., 1997. Inside the welfare contract: Discretion and accountability in state welfare administration. Social Service Review 71 (2), 1–33.

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 (4), 808–817.

Conte, M., Levy, D.T., Shahrokh, F., Staveley, J., Thompson, S., 1998. Economic determinants of income maintenance programs: The Maryland forecasting model. Journal of Policy Modeling 20 (4), 461–481.

Danielson, C., Klerman, J.A., 2008. Did welfare reform cause the caseload decline? Social Service Review 82 (4), 703–730.

Figlio, D.N., Ziliak, J.P., 1999. Welfare reform, the business cycle, and the decline in AFDC caseloads. In S.H. Danziger (Ed.) Economic conditions and welfare reform. Kalamazoo: W.E. Upjohn Institute, 17–47.

Grubb, N.W., 1984. The price of local discretion: Inequalities in welfare spending within Texas. Journal of Policy Analysis and Management 3 (3), 359–372.

Gustafsson, B., 1984. Macroeconomic performance, old age security and the rate of social assistance recipients in Sweden. European Economic Review 26 (3), 319–338.

Hainmueller, J., Hofmann, B., Krug, G., Wolf, K., 2016. Do lower caseloads improve the performance of public employment services? New evidence from German employment offices. Scandinavian Journal of Economics 118(4), 941-974.

Hill, B.C., Murray, M.N., 2008. Interactions between welfare caseloads and local labor markets. Contemporary Economic Policy 26 (4), 539–554.

Huang, C., Garfinkel, I., Waldfogel, J., 2004. Child support enforcement and welfare caseloads. Journal of Human Resources 39 (1), 108–134.

Johnson, T.R., Klepinger, D.H., Dong, F.B., 1994. Caseload impacts of welfare reform. Contemporary Economic Policy 12 (1), 89–101.

Keiser, L.R., Soss, J., 1998. With good cause: Bureaucratic discretion and the politics of child support enforcement. American Journal of Political Science 42 (4), 1133–1156.

Kimura, Y., 2006. Seikatsu hogo no kyogikai ni kakawatte (Participating in the consultative meeting on the Public Assistance and Child Rearing Allowance programs), Chiho Zaisei (Local Public Finance) 45(3), 4–10.

Klerman, J.A., Haider, S.J., 2004. A stock-flow analysis of the welfare caseloads. Journal of Human Resources 39 (4), 866–886.

Lee, M.A., Harvey, M., Neustrom, A., 2002. Local labor markets and caseload decline in Louisiana in the 1990s. Rural Sociology 67 (4), 556–577.

Lewis, W., Henry, M., 2004. Do rural welfare caseloads have a stake in metropolitan area growth? Southern Business & Economic Journal 27 (1/2), 22–36.

Lens, V., 2006. Examining the administration of work sanctions on the frontlines of the welfare system. Social Science Quarterly 87 (3), 573–590.

Lipsky, M., 1984. Bureaucratic disentitlement in social welfare programs. Social Service Review 58 (1), 3–27.

Lloyd, C., King, R., Chenoweth, L., 2002. Social work, stress and burnout: A review. Journal of Mental Health 11 (3), 255–265.

Mayer, S.E., 2000. Why welfare caseloads fluctuate: a review of research on AFDC, SSI, and the Food Stamps program. Treasury Working Paper 00/7, New Zealand Treasury.

Moffit, R., 2003. The role of nonfinancial factors in exit and entry in the TANF program. Journal of Human Resources 38 (supplement), 1221–1254.

Montiel Olea, J. L., Pflueger, C. E., 2013. A robust test for weak instruments. Journal of Business and Economic Statistics 31, 358–369.

Nakajima, M., Arakawa, T., 2004. Chihojichitai niokeru jinji-ido ni kansuru anketo chosa hokoku [Survey on the questionnaire on staff transfers in local governments]. Doshisya Policy and Management Review 5, 85–99.

Pflueger, C. E., Wang, S., 2015. A robust test for weak instruments in Stata. The Stata Journal 15(1), 216–225.

Ridzi, F., London, A.S., 2006. "It's great when people don't even have their welfare cases opened": TANF diversion as process and lesson. Review of Policy Research 23 (3), 725–743.

Schiller, B.R., 1999. State welfare-reform impacts: Content and Enforcement Effects. Contemporary Economic Policy 17 (2), 210–222.

Schiller, B.R., Brasher, C.N., 1993. Effects of workfare saturation on AFDC caseloads. Contemporary Policy Issues 11 (2), 39–49.

Schmiedera, J.F., Trenklee, S., 2020. Disincentive effects of unemployment benefits and the role of caseworkers. Journal of Public Economics 182 (2020) 104096.

Smith, B., 2005. Job retention in child welfare: Effects of perceived organizational support, supervisor support and intrinsic job value. Children and Youth Services Review 27 (2), 153–169.

Spindler, Z.A., Gilbreath, W.S., 1979. Determinants of Canadian social assistance participation rates. International Journal of Social Economics 6 (3), 164–1974.

Strolin, J.S., McCarthy, M., Caringi, J., 2007. Causes and effects of child welfare workforce turnover: Current state of knowledge and future directions. Journal of Public Child Welfare 1 (2), 29–52.

Suzuki, W., Zhou, Y., 2007. Welfare use in Japan: Trends and determinants. Journal of Income Distribution 16 (3/4), 88–109.

Takeda, F., Yokoyama, E., Miyake, T., Ohida, T., 2002. Mental health and job factors in social workers at social welfare office. Journal of Occupational Health 44 (6), 385–390.

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 (2), 225—254.

Ziliak, J.P., Figlio, D., Davis, E., Connolly, L., 2000. Accounting for the decline in AFDC caseloads: Welfare reform or economy? Journal of Human Resources 35 (3), 570–586.