

Political mergers as coalition formation: An analysis of the *Heisei* municipal amalgamations

ERIC WEESE
Yale University

In Japan, a formula-based transfer system resulted in local benefits from municipal mergers differing substantially from national benefits. A change in this transfer policy and the mergers that resulted are analyzed using a structural model involving private consumption, public good quality, and geographic distance, along with an asymmetric information problem between the national and local levels of government. The merger process is modeled using a cooperative form coalition formation game. Parameter estimates are obtained using a moment inequalities approach that requires neither an equilibrium selection assumption nor the enumeration of all possible mergers. Estimates suggest that the actual merger incentives the national government offered were weak relative to the optimal incentives, and the post-merger number of municipalities were large relative to the optimal number.

KEYWORDS. Municipal amalgamation, moment inequalities, stable set, Japan.

JEL CLASSIFICATION. D71, D82, H77.

Theoretical results such as [Alesina and Spolaore \(1997\)](#) show that political boundaries resulting from local democratic decisions may be inefficient. At the international level, there is substantial disagreement about when and how country borders can be redrawn, and at the subnational level, a variety of approaches have been employed to change local political boundaries. Currently, however, there is no empirical evidence regarding the efficiency of boundaries resulting from local decision-making. This paper analyzes a recent set of municipal mergers in Japan that relied on local approval of amalgamations, and shows that the final number of jurisdictions was around twice as large as it would have been under centralized decision-making.

Eric Weese: eric.weese@yale.edu

This is a substantially revised version of my job market paper that also incorporates work that was presented separately as “Coalition formation with panel data.” I would like to thank my thesis committee: Daron Acemoglu, Abhijit Banerjee, and Esther Duflo. I would also like to thank Nobuo Akai, Alberto Alesina, Tim Armstrong, Shun-ichiro Bessho, Thomas Fujiwara, Tal Gross, Masayoshi Hayashi, Hidehiko Ichimura, Vadim Marmer, Konrad Menzel, Masashi Nishikawa, Tai Otsu, Nancy Qian, Pablo Querubin, Bryony Reich, Testuya Shimane, Enrico Spolaore, Kota Sugahara, Francesco Trebbi, Patrick Warren, and many other attendees at seminar presentations for their helpful comments. This research was supported by a Canadian Institute for Advanced Research junior fellowship and a Japan Society for the Promotion of Science postdoctoral fellowship. Computational support was provided by the Yale Faculty of Arts and Science High Performance Computing facilities. The usual disclaimer applies.

Copyright © 2015 Eric Weese. Licensed under the [Creative Commons Attribution-NonCommercial License 3.0](https://creativecommons.org/licenses/by-nc/4.0/). Available at <http://www.qeconomics.org>.
DOI: 10.3982/QE442

In the *Heisei Daigappei*, individual Japanese municipalities could choose what merger if any they wished to participate in, given a set of national government transfer policies. Due to claimed differences in efficiencies of scale, prior to the mergers, smaller municipalities spent over ¥1,000,000 per capita per year providing services that larger municipalities provided for slightly over ¥100,000, with this difference being covered by transfers from the national government.¹ These transfers distorted local incentives: municipalities preferred to remain independent and receive large transfers in situations where the national government would rather have had them merge. A special merger-promotion transfer policy in place for a limited time led to a large number of mergers occurring within the 1999–2010 period, as shown in Figure 1. Observed mergers can thus plausibly be treated as the outcome of a single period coalition formation game, making analysis via a structural model feasible.

The theoretical model in this paper involves municipalities that provide a geographically located public good and a national government that makes equalizing transfers to

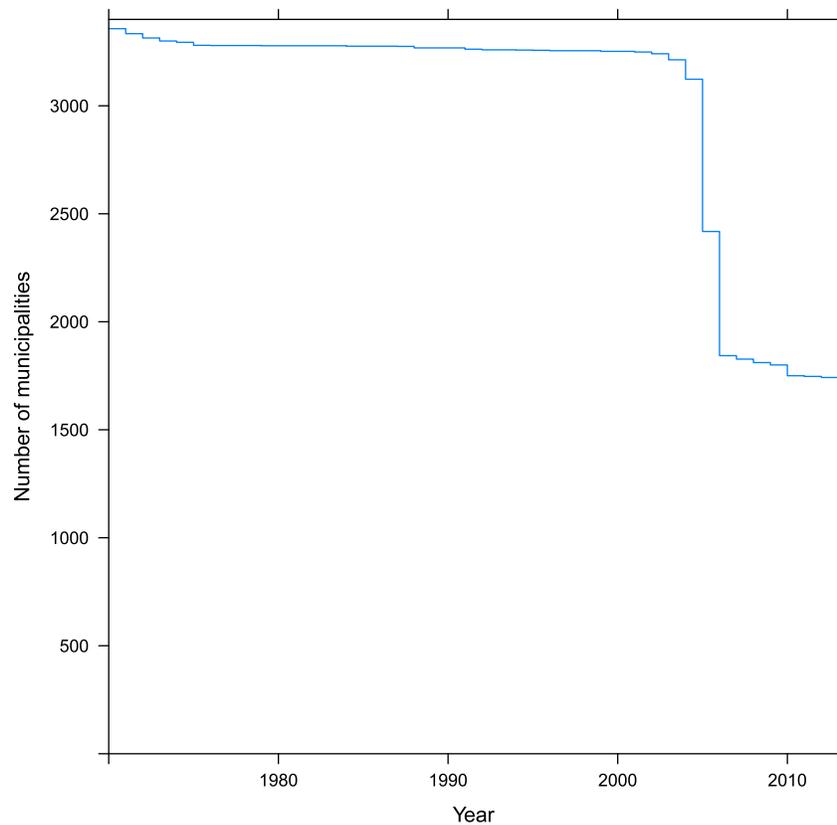


FIGURE 1. Japanese municipalities: 1970–present

¹For comparisons, ¥1 = 1¢ is a rough but useful approximation. During the period in which financial data were analyzed, the U.S. dollar/Japanese yen exchange rate varied from ¥147 = \$1 (Aug. 1998) to ¥80 = \$1 (Oct. 2010). Gross domestic product (GDP) per capita has remained relatively constant at ¥4,000,000.

those localities that have particularly high costs or low tax bases. Policy choices maximize sums of individual utilities, and for any given fixed set of borders, there is no conflict between the objectives of the local governments and that of the national government. The design of proper local borders for the federal system in this model, however, is itself subject to a classic trade-off of federalism: the national government lacks information about which potential arrangements of boundaries are idiosyncratically good or bad, while local governments do not take into account all the structural costs and benefits of potential configurations. The possibility of municipal mergers thus leads to a principal-agent problem, resulting in an optimal transfer policy different from the one the national government would offer with fixed boundaries.

The merger process is modeled as a [Bogomolnaia and Jackson \(2002\)](#) hedonic coalition formation game, where due to a commitment problem, municipalities cannot offer payments or otherwise bargain with each other in exchange for agreeing to a merger.² The solution to this game is expressed as an abstract stable set. Parameters are estimated based on a revealed preference approach, using data on actually observed mergers. Two problems that must be overcome are the very large number of potential arrangements of municipalities into mergers and the fact that coalition formation games of the type being examined do not, in general, have a unique equilibrium. The first problem is dealt with by making a specific assumption regarding the distribution of the idiosyncratic shocks, such that shocks for all potential mergers can be described by a relatively low-dimension random variable. The second problem is handled by using a moment-inequality-based estimator, following [Pakes \(2010\)](#), such that no assumption regarding the actual equilibrium selection rule is needed.

Parameter estimates based on the stability of the observed coalition structure show a trade-off between geographic proximity and efficiencies of scale in the provision of local public services. At the estimated parameters, the optimal number of municipalities from the perspective of the national government is 635 for the simplest possible model, and around 1000 using an approach based more closely on the actual municipal structure and transfer policy. In contrast, the actual final number of municipalities was 1750.

Welfare calculations based on estimated parameters are challenging because the relative weight that the national government places on different individuals is unknown. For plausible assumptions, if the national government had decided to mandate a pattern of mergers, the benefit from these centralized mergers is equivalent to between ¥520 billion and ¥912 billion in public funds, compared to the case with no mergers. Estimates of the benefits of the decentralized mergers that actually occurred are even more challenging, as simulation is computationally infeasible. An extremely rough estimate gives total benefits about the same as those of the centralized mergers, with half of this due to the idiosyncratic term.

At the estimated parameters, the merger incentives actually offered by the national government appear substantially weaker than the optimal incentives. Simulations in-

²[Acemoglu \(2003\)](#) argues that the case where transfers are not possible is more relevant for issues in political economy. The approach in this paper is thus necessarily different from that used in [Baccara, Imrohoroğlu, Wilson, and Yariv \(2012\)](#), which is based on [Fox \(2010\)](#). The hedonic coalition formation model also presents different issues than those encountered in two-sided matching games ([Roth \(2008\)](#)).

volving pairwise mergers suggest that these weak incentives could account for the difference between the optimal pattern of amalgamations and the actual outcome.

The main conclusion of the paper is thus that there was a substantial benefit from mergers but there were too few of them, and this problem can be explained by incentives that were too weak. Even with these suboptimal incentives, however, it is not clear that centralized mergers would have been a superior option. The major contribution is thus to the literature on optimal jurisdiction boundaries, and, in particular, models of the sort presented in [Alesina and Spolaore \(2003\)](#).³ Theoretical inefficiency results for these models are well known.⁴ The first empirical evidence that this sort of inefficiency is economically important, however, is provided in this paper.

The empirical approach differs substantially from papers such as [Alesina, Baqir, and Hoxby \(2004\)](#), which shows that heterogeneity affects the size of school districts and other local jurisdictions, and [Alesina, Baqir, and Easterly \(1999\)](#), which shows that the heterogeneity of jurisdictions affects the choice of public goods. The reduced-form approach employed in these papers is well suited to studying the effects that they focus on, but it is not simple to extend this approach to determine, for example, the optimal number of jurisdictions.⁵ A structural model is more appropriate in this case, and basing this model on a cooperative form coalition formation game is particularly attractive because it does not require knowledge of all the details of the actual process by which municipalities merged. Estimation without a fully specified data generating process is possible because moment inequality estimation requires only that some outcomes are (on average) ruled out: as this sort of ruling-out is precisely what much of cooperative game theory focuses on, the two techniques are well suited for use together.

From the perspective of estimation, the closest related work is [Gordon and Knight \(2009\)](#), who examine school district mergers in Iowa.⁶ In their model, districts merge in pairs and match quality is symmetric. This results in a unique stable matching, and parameters are estimated via simulated method of moments. The approach used in this paper, on the other hand, allows for more than two partners, but does not guarantee a unique equilibrium. The two approaches are thus complementary: the model presented below is applicable to more cases, while the model used by [Gordon and Knight](#) has desirable properties such as uniqueness.

³Earlier models involving coalition formation with heterogeneous preferences include [Greenberg and Weber \(1986\)](#) and [Demange \(1994\)](#).

⁴For example, inefficiency is discussed in all eight of the theory papers compared in [De Donder, Le Breton, and Peluso \(2012\)](#).

⁵Specifically, the optimal number of jurisdictions in an [Alesina and Spolaore \(1997\)](#) style model depends on the “model” population density (relative to preference heterogeneity), which does not map directly to any population density in the data. The simplest solution to this issue appears to be to expand the model from one to two dimensions and to use geographic population density, which is precisely what is done below. The stability arguments in [Alesina and Spolaore \(1997\)](#), which implicitly specify a coalition formation game, do not extend easily to two dimensions. The detailed treatment in [Haimanko, Le Breton, and Weber \(2004\)](#) similarly appears not to extend to the two dimensional case, and thus a more general coalition formation game needs to be used. These games have the potential for multiple equilibria, and it is unclear how this could be accounted for in a reduced-form estimation strategy.

⁶[Desmet, Le Breton, Ortuño-Ortín, and Weber \(2011\)](#) use a theoretical model closer to this paper, but fit the model via calibration.

Section 1 presents a model of local public goods and municipal mergers. Section 2 discusses the Japanese data and, in particular, the transfer policy actually used. Section 3 describes the estimation strategy, and Section 4 presents the parameter estimates and uses them to compare the mergers that actually occurred to those that the central government would have chosen. Section 5 extends the model by characterizing the optimal merger incentives for the national government to offer local governments, and comparing these to the incentives actually offered. Appendices A and B follow the conclusion, Section 6. Appendices C–F and additional material are available in supplementary files on the journal website, <http://qeconomics.org/supp/442/supplement.pdf> and http://qeconomics.org/supp/442/code_and_data.zip.

1. THEORY

1.1 Local governments

There is a single country, populated by individuals who are distributed across a plane. The location of these individuals is fixed and they are partitioned into municipalities. Each municipality $m \in M$ provides a public good of quality q_m to its N_m residents at a single location θ_m on this plane. Providing this good costs $q_m c(X_m)$, where the cost $c(X_m)$ of providing one quality unit of the good depends on the covariates X_m of the municipality. The case where X_m is scalar is particularly easy to analyze: below, cost will be expressed as $c(N_m)$ in examples where total population is assumed to be the only factor determining cost.

Municipality m levies taxes at rate τ_m on tax base $Y_m = \sum_{i \in m} y_i$, where i indexes individuals.⁷ There is also a transfer from the national government: municipality m receives T_m regardless of the quality of service it chooses to provide. Feasible (q_m, τ_m) pairs are determined by the municipal budget constraint

$$q_m c(X_m) = \tau_m Y_m + T_m. \quad (1)$$

Individual utility is assumed to take the additively separable form

$$u_i(q_m, \tau_m, \theta_m) = \beta_0 \log((1 - \tau_m)y_i) + \beta_1 \log(q_m - \beta_3) + \beta_2 \ell(i, \theta_m) + \epsilon_m, \quad (2)$$

where β_3 is some minimum level of public good provision, and $\ell(i, \theta)$ is the distance between the location of individual i on the plane and the location θ of the public good provided by the municipality of which i is a member.⁸ The ϵ_m term is an idiosyncratic shock and will be assumed to be normally distributed with mean zero. As a scale normalization, the variance of this shock will be assumed to be 1. Each ϵ , while independent of covariates X , may be correlated with other shocks: this correlation structure will be discussed in Section 3.

⁷Treating m as a set rather than an index number is slightly nonstandard, but eliminates the need for additional notation when discussing estimation in Section 3.

⁸A restriction imposed throughout this paper is that individuals do not move or otherwise change their ideal point. A similar model could be constructed, however, with each individual owning a home, the value of the home varying with distance to the public good, and people voting to maximize their real estate value.

The first two terms of the utility function in (2) have Stone–Geary form, with the minimum level of the private good set to zero. As these are the only terms that contain q_m or τ_m , all individuals will share the same ideal point τ_m^* for taxation. To see this, note that (2) can be rewritten to treat income as an individual fixed effect,

$$u_i(q_m, \tau_m, \theta_m) = \beta_0 \log(1 - \tau_m) + \beta_1 \log(q_m - \beta_3) + \beta_2 \ell(i, \theta_m) + \alpha_i + \epsilon_m, \quad (3)$$

where $\alpha_i = \beta_0 \log(y_i)$. On the other hand, there is no agreement among individuals regarding the location θ_m at which the public good should be provided. The set of feasible points is a plane; thus choosing θ_m^* is a multidimensional policy decision, a problem that has no single accepted solution concept.

To resolve this, political decision-making is assumed to take place in a probabilistic voting framework, with the standard result that the selected policy maximizes a weighted sum of individual utilities. This is discussed in more detail in Appendix C. At the local level, assume that equal weight is given to all individuals: the local politician acts as a Benthamite social planner. Let $u_{mm}(T_m)$ correspond to the maximum of the local politician's objective function in municipality m , given transfers T_m ,⁹

$$u_{mm}(T_m) = \beta_0 \log(1 - \tau_m^*) + \beta_1 \log(q_m^* - \beta_3) + \beta_2 \ell_m(\theta_m^*) + \alpha_m + \epsilon_m, \quad (4)$$

where the average fixed effect is $\alpha_m = \frac{1}{N_m} \sum_{i \in m} \alpha_i$, the average distance is $\ell_m(\theta) = \frac{1}{N_m} \sum_{i \in m} \ell(i, \theta)$, and the optimal policies for τ , q , and θ are given by

$$\begin{aligned} \tau_m^* &= 1 - \frac{\beta_0}{\beta_0 + \beta_1} \frac{Y_m + T_m - \beta_3 c(X_m)}{Y_m}, \\ q_m^* &= \frac{\beta_1}{\beta_0 + \beta_1} \frac{Y_m + T_m - \beta_3 c(X_m)}{c(X_m)} + \beta_3, \\ \theta_m^* &= \operatorname{argmin}_{\theta} \ell_m(\theta). \end{aligned} \quad (5)$$

Now consider the possibility of municipal mergers. Let $S \subset M$ be a coalition of municipalities that will merge together.¹⁰ For each coalition S , the municipalities in S are permanently eliminated and a single new amalgamated municipality is created. The set of all possible partitions π of municipalities into coalitions will be represented by Π .

An amalgamated municipality behaves exactly as outlined above, and is not involved in any further mergers. That is, the utility for individual i in merger S will be as in (3), except replacing m with S :

$$u_i(\tau_S, q_S, \theta_S) = \beta_0 \log(1 - \tau_S) + \beta_1 \log(q_S - \beta_3) + \beta_2 \ell(i, \theta_S) + \alpha_i + \epsilon_S. \quad (6)$$

Assume that, while the national government commits in advance to making a transfer T_S to the amalgamated municipality, it is not possible for the municipalities in S to commit to a given τ_S , q_S , or θ_S in advance of the merger. Further assume that the sufficient conditions for a unique (q^*, τ^*, θ^*) political equilibrium, described in Appendix C,

⁹The second m subscript will be used to indicate mergers.

¹⁰Notation used follows Banerjee, Konishi, and Sönmez (2001) to the extent possible.

hold for each merger. The post-merger choice of q_S^* , τ_S^* , and θ_S^* is then known in advance for each potential coalition S , and is exactly as in (5), except replacing municipality m with coalition S .

The payoff for municipality m participating in coalition S can be expressed in the same form as (4):¹¹

$$u_{mS} = \beta_0 \log(1 - \tau_S^*) + \beta_1 \log(q_S^* - \beta_3) + \beta_2 \ell_m(\theta_S^*) + \alpha_m + \epsilon_S. \quad (7)$$

Municipal mergers are thus a pure hedonic coalition formation game, where the payoff to each player depends only on the coalition to which (s)he belongs and not on what other coalitions occur.¹² The third term of (7), related to the distance on the plane to the location of the public service, distinguishes the setup from that of [Gordon and Knight \(2009\)](#). This is because this term differs across municipalities in S , and thus municipality m may prefer merger S to S' even if municipality m' has the opposite preference.¹³

Now consider the difference in payoffs for municipality m of participating in coalition S versus remaining a singleton:¹⁴

$$\begin{aligned} u_{mS} - u_{mm} &= \beta_0 (\log(1 - \tau_S^*) - \log(1 - \tau_m^*)) \\ &\quad + \beta_1 (\log(q_S^* - \beta_3) - \log(q_m^* - \beta_3)) \\ &\quad + \beta_2 (\ell_m(\theta_S^*) - \ell_m(\theta_m^*)) \\ &\quad + \epsilon_S - \epsilon_m. \end{aligned} \quad (8)$$

Here the payoffs to municipality m depend only on the characteristics of m and S , and a transfer policy to which the national government has already committed. The first term is positive if the tax rate will decrease, which is the case when $\frac{Y_S + T_S - \beta_3 c(X_S)}{Y_S} > \frac{Y_m + T_m - \beta_3 c(X_m)}{Y_m}$. The second term is positive if the quality of the public good provided will be higher, which is the case when $\frac{Y_S + T_S - \beta_3 c(X_S)}{c(X_S)} > \frac{Y_m + T_m - \beta_3 c(X_m)}{c(X_m)}$.¹⁵ These two terms are the channel through which the national government's transfer policy affects municipal merger decisions: if T_S is higher or T_m is lower, municipality m is more likely to prefer the merger over remaining a singleton.

¹¹The first term here is the utility received by individuals in municipality m from their private consumption; however, due to the log functional form, Y_m moves into the fixed effect α_m in the same way as in (3). The first term will thus be the same for all municipalities participating in coalition S , as the tax rate is by assumption the same throughout a given amalgamated municipality. Similarly, the level of public goods is also assumed to be the same within the same amalgamated municipality. Thus, the second term will also be the same for all municipalities participating in coalition S , as all residents are assumed to value public goods equally.

¹²This is the game introduced by [Dreze and Greenberg \(1980\)](#), except without the possibility of even within-coalition transfers.

¹³Preference alignment is required for coalition formation games to have a solution of the form presented in [Farrell and Scotchmer \(1988\)](#).

¹⁴For simplicity, the payoff to the singleton merger $\{m\}$ is denoted by u_{mm} , following (4), rather than $u_{m\{m\}}$.

¹⁵The third term will always be zero or negative as long as distance is undesirable: θ_m^* was the minimizer of ℓ_m and, thus, θ_S^* cannot be better than θ_m^* for municipality m .

1.2 National government

The national government offers transfers T to municipalities. Funds for these transfers come from an outside source, and the government spends enough of these on activities that are outside of the model that the marginal opportunity cost of providing transfers can be treated as constant. Let this marginal cost of funds for the national government be b . Political decision-making at the national level also takes place in a probabilistic voting framework, here with possibly unequal weights across individuals in different municipalities.¹⁶

If there is no possibility of mergers, the national government's objective is

$$W(T) = \sum_{m \in M} w_m u_{mm}(T_m) - b \sum_{m \in M} T_m, \quad (9)$$

where w_m is the weight placed on municipality m . For each municipality m the τ_m^* , q_m^* , and θ_m^* policies chosen by the local government are exactly the policies the national government would want the municipality to select. Optimal transfers for the national government in this "fixed boundary" case are determined by the first order conditions and are

$$T_m^{\text{FB}} = \beta_3 c(X_m) - Y_m + \frac{w_m(\beta_0 + \beta_1)}{b}. \quad (10)$$

In the special case where all individuals have the same y and are weighted equally by the national government, (10) further simplifies to

$$T_m^{\text{FB}} = \beta_3 c(X_m) - a Y_m, \quad (11)$$

where $a = 1 - \frac{w}{yb}(\beta_0 + \beta_1)$. An important feature of this optimal transfer scheme is that it provides higher transfers to municipalities that are facing higher costs. Intuitively, this will cause difficulties when there is the possibility of municipal mergers that would lower the per capita cost of services: the benefit to the national government of such a merger will be greater than the benefit of that merger to the municipalities themselves.

Section 5 will describe the optimal transfer policy when there is the possibility of mergers. Up to that point, however, it is only necessary to have the above characterization of the optimal transfer policy with a given set of fixed boundaries: the merger incentives discussed in Section 2 and used in estimation in Sections 3 and 4 can simply be treated as given and arbitrary deviations from the fixed boundary optimum policy.

1.3 Solution concept

If there are more than two municipalities, situations can arise where the pattern of mergers that will emerge is unclear. In particular, preference cycles can occur, as in the classic "roommates problem."

¹⁶The transfers considered here differ slightly from those in Haimanko, Le Breton, and Weber (2005) and related papers, as they are offered to jurisdictions by a higher level of government, rather than being offered by jurisdictions to individual residents.

EXAMPLE 1 (Gale and Shapley (1962)). Suppose $M = \{1, 2, 3\}$ and preferences are

$$\begin{aligned} \{1, 2, 3\} <_1 \{1\} <_1 \{1, 3\} <_1 \{1, 2\}, \\ \{1, 2, 3\} <_2 \{2\} <_2 \{1, 2\} <_2 \{2, 3\}, \\ \{1, 2, 3\} <_3 \{3\} <_3 \{2, 3\} <_3 \{1, 3\}. \end{aligned} \quad (12)$$

Given these preferences, there is no immediately obvious solution to this coalition formation game: players 2 and 3 would both like to deviate from the $\{\{1, 2\}, \{3\}\}$ partition, and there are similar deviations for other partitions.¹⁷

To resolve this issue, Ray and Vohra (1997) develop a solution concept based on only considering refinements: deviations that involve a subset of a single coalition, and thus result in moving from a coarser partition to a finer one. They then define the coarsest partitions that do not have any refinements as the solution to the coalition formation game. In the roommates example, then, the $\{\{1, 2\}, \{3\}\}$ partition would be considered stable, as the $\{2, 3\}$ coalition is not a subset of a coalition in the partition. There would thus be three solutions to Example 1.

The environment considered in this paper is simpler than that considered in Ray and Vohra (1997), and thus a simpler solution concept allowing both splits and mergers of coalitions can be used.

THEOREM 1. *Let Π^* be the set of partitions that do not have any deviations that are refinements or coarsenings. Then Π^* exists, is not empty, and is unique.*

See Appendix A for the proof.

Although Π^* is unique, it may contain multiple partitions. Multiplicity of solutions is a fundamental property of roommate-type coalition formation games (Barberà and Gerber (2007)). Intuitively, as Example 1 is symmetric, any plausible solution concept that gives $\{\{1, 2\}, 3\}$ should also give the partition with $\{2, 3\}$ and the partition with $\{1, 3\}$ as solutions as well.

The interpretation of Π^* is as follows. If $\pi \notin \Pi^*$, then there is some coalition that would definitely deviate from π , and thus π should not be observed as the outcome of the coalition formation game. On the other hand, if $\pi \in \Pi^*$, then π might be observed as the outcome of the game. The solution concept used thus rules out partitions that should definitely not occur, but does not specify precisely what partition will occur.¹⁸

For the estimation approach used in Section 3, the following restrictions implied by Theorem 1 will be used, given a $\pi \in \Pi^*$:

$$\forall S' \subset S \in \pi, \exists m \in S' \quad \text{s.t.} \quad u_{mS} > u_{mS'}, \quad (13)$$

$$\text{for } \{m\}, \{m'\} \in \pi, \text{ and } S' = \{m, m'\}, \quad \text{either } u_{mm} > u_{mS'} \text{ or } u_{m'm'} > u_{m'S'}. \quad (14)$$

¹⁷Bogomolnaia, Le Breton, Savvateev, and Weber (2008) show that the core can be empty in a standard jurisdiction formation model, and existence appears to require strong restrictions on the population distribution even when individuals are placed on a line.

¹⁸The approach used by Gordon and Knight (2009) to guarantee uniqueness is not possible in this case because the $\ell(i, \theta_m)$ in (3) generates the possibility of preference cycles as in Example 1. The noncooperative approach used by Diermeier (2003) requires additional assumptions regarding the exact process by which coalitions were formed.

The first restriction states that for any subcoalition S' of a coalition in the partition π , at least one member of this subcoalition must prefer his/her current coalition to the subcoalition. The second restriction states that for any two singletons in the partition π , at least one of these players must not have wanted to merge with the other.¹⁹

2. JAPANESE CONTEXT

Japan is a unitary state divided into 47 prefectures, whose boundaries have remained roughly unchanged since the 1890s. As shown in Figures 2 and 3, each of these prefectures is divided into municipalities. Municipalities are responsible for providing public services in six major areas: firefighting, public works, education, welfare, industry, and

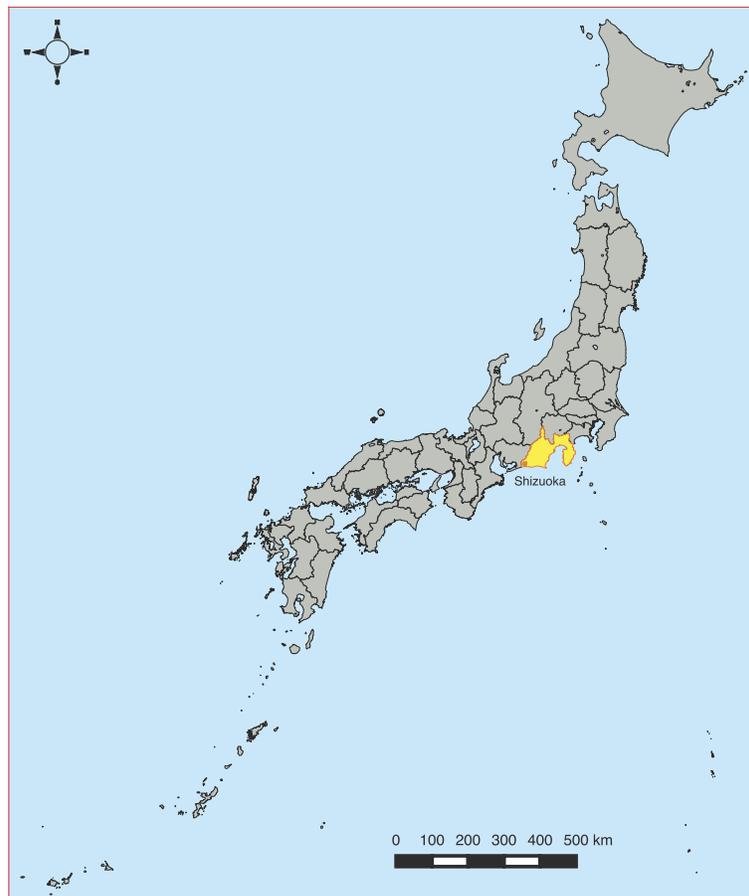


FIGURE 2. Prefectures of Japan.

¹⁹The restrictions implied by Theorem 1 for coarsenings involving nonsingleton coalitions are computationally difficult because of the recursion involved in considering deviations in the Ray and Vohra style setup. Larger coarsenings are thus not used in Section 3.

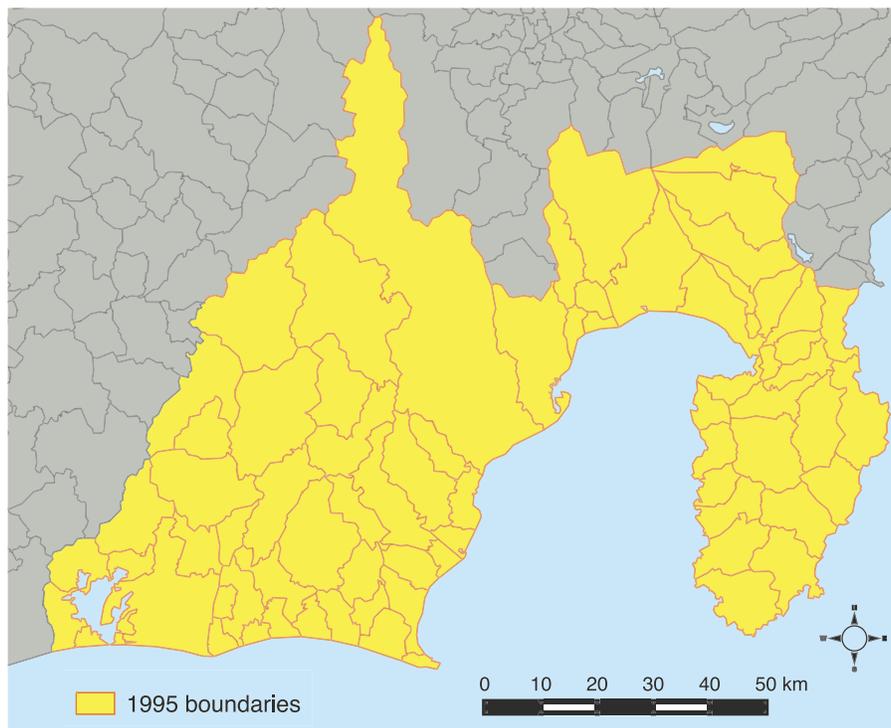


FIGURE 3. Shizuoka prefecture (pre-merger).

administration.²⁰ To analyze the municipal mergers that occurred in Japan using the theoretical model presented above, it is necessary to make simplifying assumptions so as to express the data in a form that matches (1) (the municipal budget constraint) and (3) (the individual utility function).

The municipal budget constraint involves receipt of lump sum transfers from the national government, expenditures on public good provision, and taxation of the municipal tax base. The transfer system is central to this paper, and as it is closely connected to the cost of providing public goods, these two components of the budget constraint are discussed at some length below. The third component—local taxation—is calculated directly from published tax base data as described in Appendix D.1. One issue with taxation is that there is some bureaucratic imprecision regarding the $\bar{\tau}Y_m$ term in (15). While *de jure* municipalities are allowed to set their own tax rates, *de facto* it appears that rates less than $\bar{\tau}$ may be prohibited. In Section 3, parameters will be estimated both for the case where $\bar{\tau}$ acts as a minimum tax rate and where τ^* can be chosen freely.

The individual utility function includes terms from the municipal budget constraint as well as the idiosyncratic shock ϵ and distance ℓ .²¹ The idiosyncratic term will be in-

²⁰Public goods that could generate substantial externalities appear to be provided by higher levels of government, rather than by municipalities. For example, waterways and major roads are the responsibility of prefectures.

²¹The individual fixed effect α_i is not relevant because the data used involve choices between coalitions; thus this term will always disappear, as shown in (8).

terpreted as cultural heterogeneity and will be discussed further in Section 3. Distance is taken to be geographic distance and is calculated using census grid square data on population. This is appropriate because the majority of local public services are provided at physical facilities, such as schools, nursing homes, libraries, and city hall itself. Further discussion and details of the distance calculations are provided in Appendix D.2.

The remainder of this section explains the national government transfer policy: first the policy when there are no mergers and then the specific merger incentives that were put in place. A major part of the transfer policy is an official report regarding the cost of providing public goods; thus the cost function c is also discussed below.

2.1 Fixed-boundary transfer policy

Beginning in the 1950s, the Japanese government established a “national minimum” reference quality for local government services. To ensure that every municipality had sufficient funds to offer services above this minimum level, the national government developed a complicated system of transfers, now called the *local allocation tax*.²² Transfers to municipalities were determined by a formula quite similar to (11):

$$T_m = \max(\tilde{c}(X_m) - 0.75\bar{\tau}Y_m, 0). \quad (15)$$

Here \tilde{c} calculates the *standard fiscal need* of a municipality: this is the level of spending that the national government reports is necessary for a municipality with characteristics X_m to provide the national minimum level of service. To determine transfers, this amount is compared to the tax revenue $\bar{\tau}Y_m$ that the municipality should be able to collect if taxes are charged at the “standard” rate. As will be discussed below, the national government changed \tilde{c} as part of its merger incentive policy. Let \tilde{c}^0 be the function in use prior to the merger period.

The relationship between the national government’s report \tilde{c}^0 and the true cost c of providing the public good has been discussed at some length in the domestic Japanese literature. There is a general impression that \tilde{c}^0 overstated cost for smaller municipalities. If true, then the efficiencies of scale implied by \tilde{c}^0 were more substantial than those actually exhibited by c , and smaller municipalities would have been receiving larger transfers than if (15) used the true c .²³

Interpreted in the context of the model from Section 1, this situation corresponds to higher weights w_m for smaller municipalities, in per capita terms. These higher weights would then result in larger optimal transfers for those municipalities in (10), relative to the case where all individuals were weighted equally. The following assumption will thus

²²The slightly confusing name is due to the fact that it is an *allocation* to local governments from *taxes* collected by the national government.

²³One potential reason why the national government might have exaggerated efficiencies of scale in this manner is that rural areas are overrepresented in the National Diet, and these areas have municipalities with much smaller populations than the underrepresented urban centers. Horiuchi and Saito (2003) provide reduced-form evidence that transfers are indeed related to malapportionment. If there were a political desire to make larger transfers to rural areas, reporting a \tilde{c} different from c could be a convenient way for the national government to effect these transfers without attracting unwanted attention.

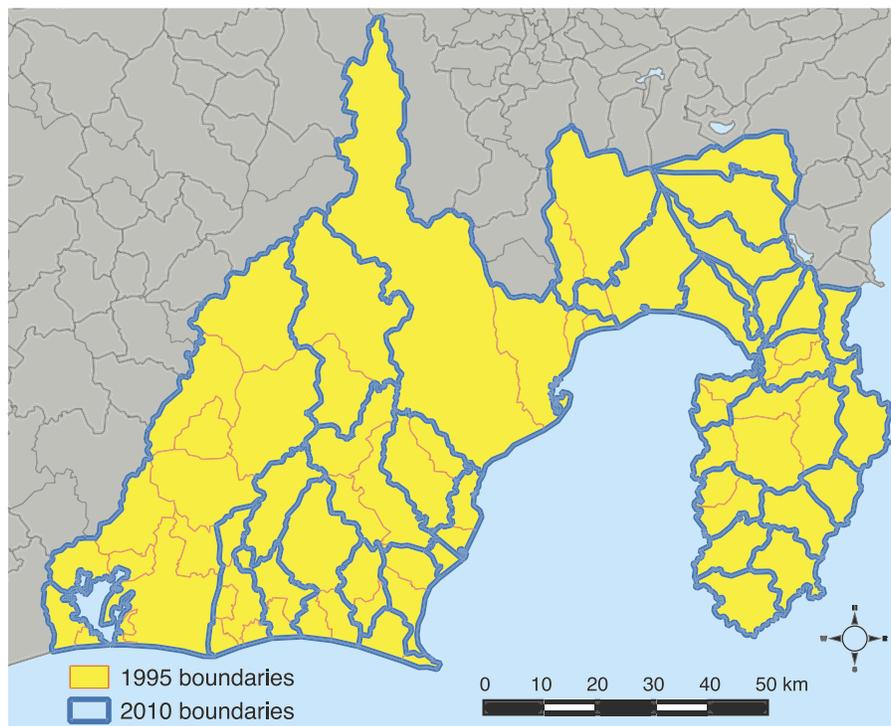


FIGURE 4. Shizuoka prefecture (post-merger).

be used to match the observed pre-merger period transfers to the theoretical model: both the national government and the local governments are aware of the true cost function c , but the national government chooses to use the (potentially different) cost function \tilde{c}^0 in its published calculations so as to produce a T^0 via (15) that maximizes its objective given its w weight vector.²⁴

In the data, the actual calculation the national government used for \tilde{c}^0 involved a sum over many splines, some of which are themselves the products of other splines, each of which depends on X_m in a different way. The end result of this calculation is shown in Figure 7. The specifics of the procedure used are described in Appendix B.

In general the calculation of \tilde{c} involves components that can roughly be divided into two groups. The first group consists of specific, itemized inputs for specific public services: for example, personnel costs associated with fire fighters. The items in this group have remained largely unchanged since the 1960s, and the national government's report of how much expenditure is necessary for fire fighters does not vary radically from year to year. There is a second group of components of \tilde{c} , however, that appear and disappear with some frequency and involve amounts that can fluctuate dramatically.²⁵ The

²⁴The more restrictive case where the national government must use $\tilde{c}^0 = c$ will be used as a robustness check.

²⁵These are labeled as either (un-itemized) capital expenditures or with vague terms such as "development" or "miscellaneous."

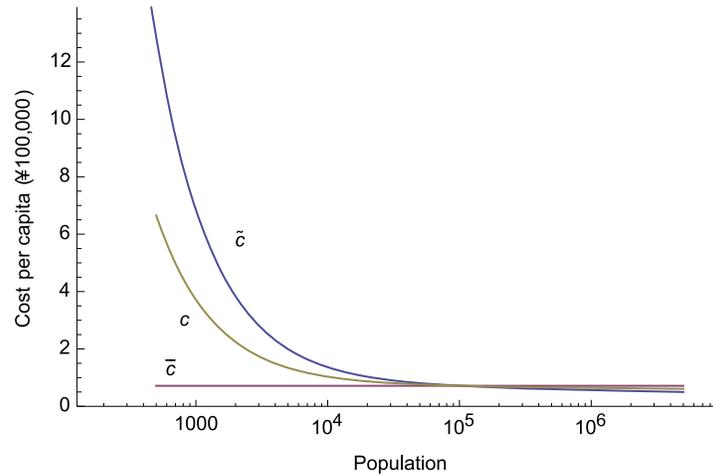


FIGURE 5. True and reported cost.

hypothesis in this paper is that the efficiencies of scale associated with this second group of expenditures are spurious, and account for the difference between the reported \tilde{c} and the true c .²⁶

For the purpose of estimation, suppose that the relationship between c and \tilde{c} can be described by a scalar ψ . The calculation of \tilde{c} described in Appendix B will be used in the actual estimator, but for illustration consider the simplified cost function shown in Figure 5:

$$\begin{aligned}\tilde{c}^0(N_m) &= (1 + \tilde{H}^0(N_m))\tilde{c}, \\ \tilde{H}^0(N_m) &= \psi H(N_m).\end{aligned}\tag{16}$$

Thus, the national government distorts the true cost function when it chooses a ψ different than 1.²⁷ The hypothesis discussed above—that the efficiencies of scale associated with the second group of expenditures in fact do not exist—corresponds to a ψ of about 0.5 and is shown in Figure 6. For estimation, \tilde{H}^0 is known from government documents, while H is unknown. Thus, the parameterization $H = \beta_4 \tilde{H}^0$ will be used in Sections 3 and 4.²⁸ The estimates presented in Section 4 give a value of $\hat{\beta}_4$ that does indeed roughly correspond to the situation in Figure 6.

²⁶Specifically, the second group is taken as those components that were later grouped into the *houkatsu* (general) component. This grouping is not related to (40), which separates adjustment coefficients based on whether they are truly adjusting a unit cost.

²⁷This is equivalent to the national government reporting $\tilde{c}(N_m) = (\psi_0 + \psi_1 H(N_m))\tilde{c}$. As there are no natural units for public good quality, normalize $\psi_0 = 1$. This normalization of the cost function then determines the units for β_1 and β_3 in the individual utility function of (3).

²⁸One hypothesis of particular interest is that there are in reality no efficiencies of scale, but this corresponds to $\psi \rightarrow \infty$ and is thus difficult to test using the parameterization involving ψ . Using β_4 , the case where there are no efficiencies of scale in the production of public goods corresponds to $\beta_4 = 0$, and the case where the national government's reported cost \tilde{c} is correct corresponds to $\beta_4 = 1$.

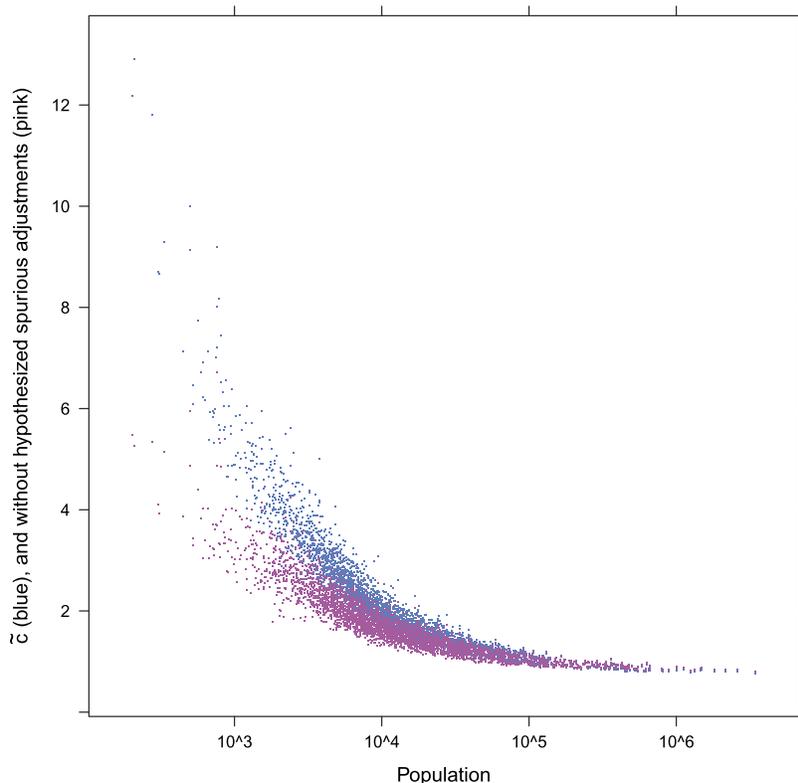


FIGURE 6. Reported and hypothesized true costs.

2.2 Merger policy

If a coalition S of municipalities engaged in a merger, then the transfers for the amalgamated municipality were calculated as

$$T_S = \underline{T}_S + \Delta_S, \quad (17)$$

$$\underline{T}_S = \max(\tilde{c}(X_S) - 0.75\bar{\tau}Y_S, 0).$$

Here \underline{T}_S is the “base” amount, i.e., what the transfer would be following (15) and treating S as a single municipality. An additional amount Δ_S was also provided: prior to the merger period, the Δ_S^0 provided was small, but due to policy changes, the merger period Δ_S^1 was quite substantial.

Figure 7 shows that, in per capita terms, \tilde{c}^0 was declining in population. Thus, prior to the merger period, any mergers that did happen to occur would result in substantial savings for the national government, because the post-merger transfer would be lower than the total pre-merger transfer. As noted in Section 1, this leads to local governments opposing mergers that the national government would like to see occur. Figure 1 shows that very few mergers occurred during this period.

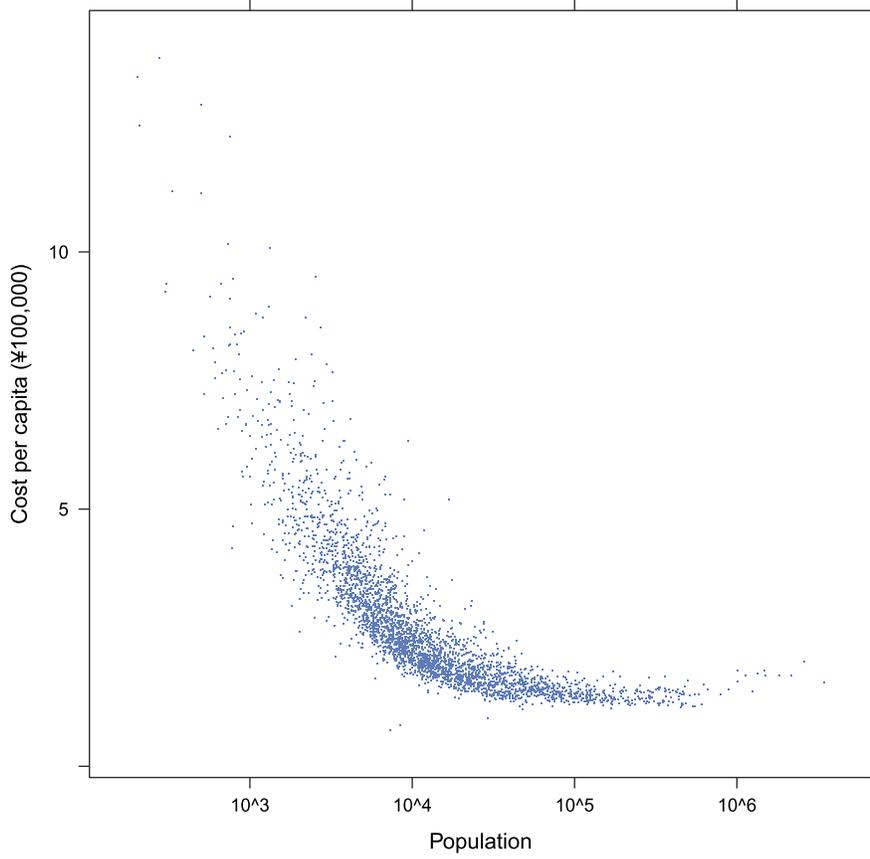


FIGURE 7. Cost of public good (according to national government).

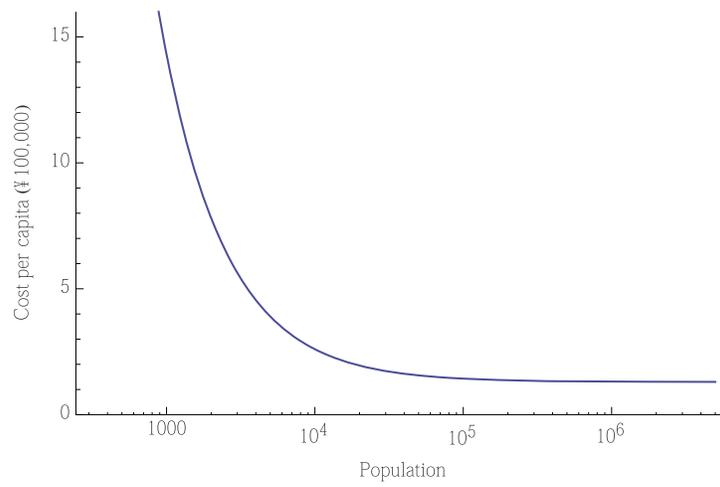


FIGURE 8. Cost of public good (Section 5).

Beginning around 2000, however, the national government introduced special transfer policies that provided an incentive for municipalities to merge: these policies included both a “stick” and a “carrot.” The details of these policy changes are provided in Appendix B.1. A simplified summary is that transfers were reduced for all municipalities, with more severe reductions for small population municipalities. Municipalities that merged, however, would be able to issue special bonds (*tokureisai*) that were very heavily subsidized by the central government. In addition, merging municipalities benefited from a special calculation (the *santeigae*) that ensured that, for the decade following a merger, the total transfer a coalition received would not decrease as a result of the merger.

Suggestive evidence for the effect of the first of these policies (the stick) on merger activity can be seen by looking at how the fraction of municipalities merging varies at different levels of cuts to transfers.²⁹ Figure 13 shows that municipalities facing a stronger incentive were more likely to participate in mergers. This effect is frequently noted in the domestic Japanese literature (e.g., Hirota (2007)), and is usually analyzed by looking at the relationship between population (highly correlated with the stick) and merger activity: Figure 14 shows this relationship.

Another effect frequently discussed in the Japanese literature is that poorer municipalities were more likely to merge (e.g., Kido (2008), Hirota (2012)). This is expected given the model in Section 1: the same size incentive corresponds to a larger change in the tax rate when the tax base is smaller, and this effect is amplified by the minimum quality β_3 for the public good. Figure 15 shows that municipalities with a lower (per capita) tax base are more likely to participate in mergers. The standard in the Japanese literature is to consider the “*fiscal strength* index” (tax base as a fraction of cost): this is shown in Figure 16.³⁰

3. ESTIMATION

Parameter estimates $\hat{\beta}$ will be obtained by examining the mergers that actually occurred in Japan and comparing them to mergers that could have occurred but did not. Municipal mergers did not cross prefectural boundaries, and thus the outcome of the coalition formation game in each prefecture is a partition of the municipalities in that prefecture.³¹ For any potential coalition S , the tax rate and public good quality can be calculated for given parameters β by replacing m with S in (5).³² Geographic distance can be

²⁹The strength of the carrot incentives depended on the exact membership of the coalition that formed; combinatorial issues make it difficult to show the effect of these incentives in a similar fashion to the stick incentives.

³⁰Remote islands (*ritou*) are excluded from the calculations for Figures 13–16, as in this case geographic considerations are particularly important.

³¹There is one exception, involving a single municipality switching prefectures. It is treated as though the municipality in question was always part of the “destination” prefecture.

³²The parameter β is five dimensional. There are four parameters of interest from (3): the value of private consumption (β_0), the value of public consumption (β_1), the disutility of distance (β_2), and the minimum quality for the public good (β_3). The degree to which the true efficiencies of scale differ from the national government’s reported function (β_4) enters via the cost function c . Estimation of models with a larger num-

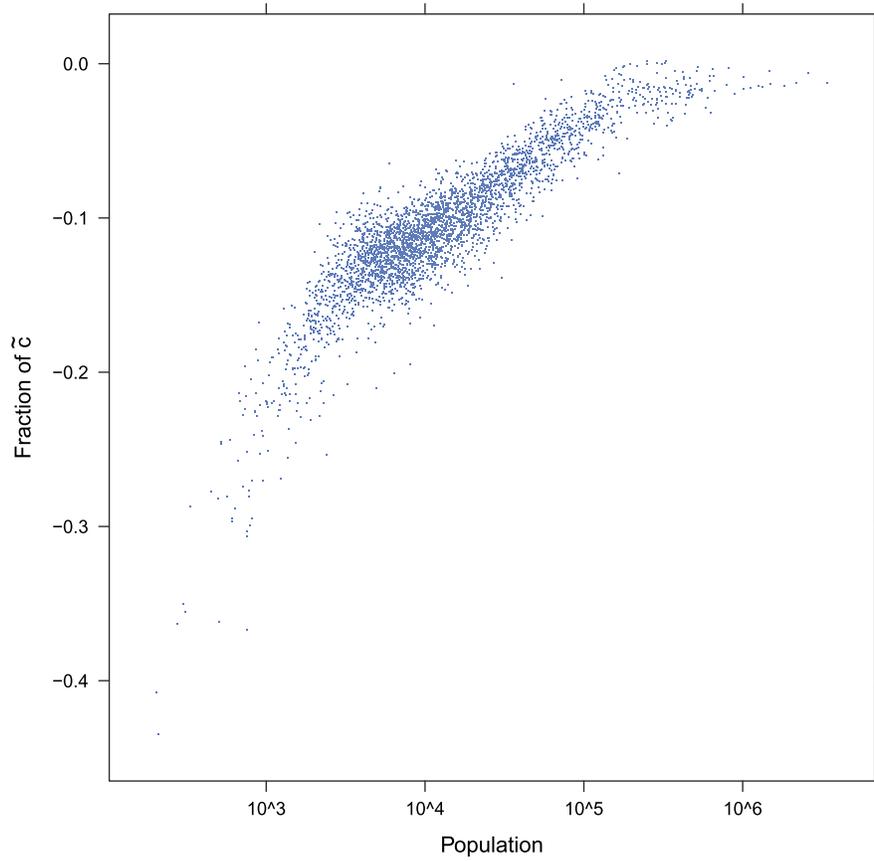


FIGURE 9. Stick (actual policy).

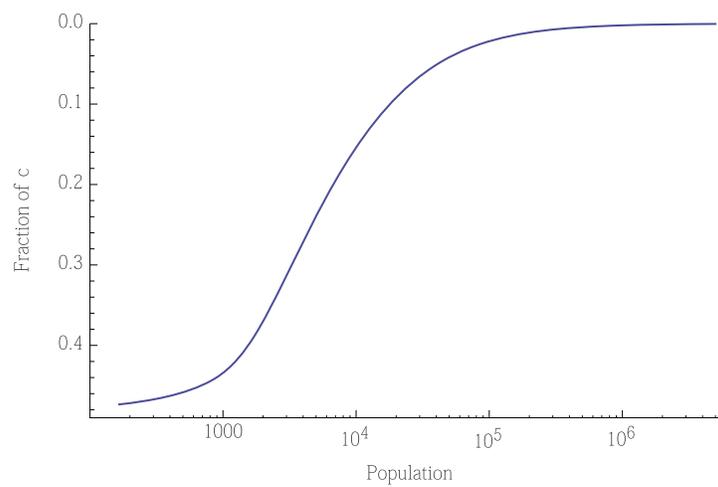


FIGURE 10. Stick (Section 5).

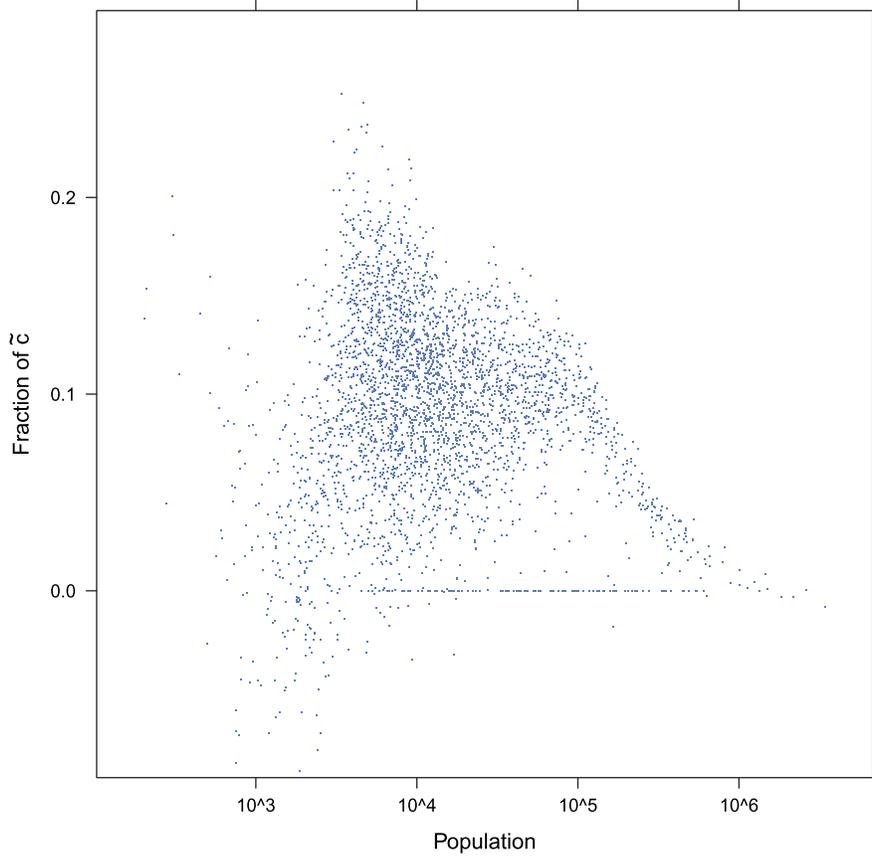


FIGURE 11. Carrot (actual temporary policy treated as permanent).

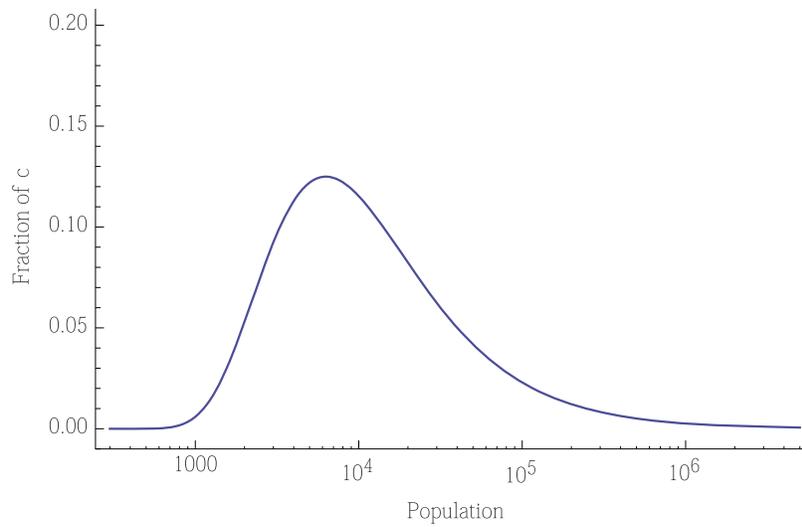


FIGURE 12. Carrot (Section 5).

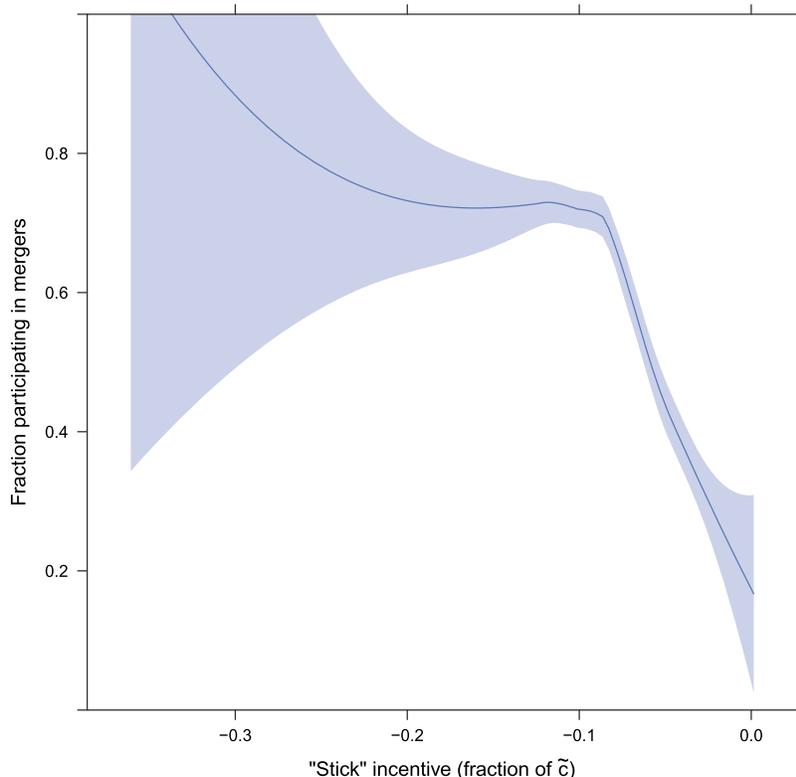


FIGURE 13. Mergers by merger incentive.

calculated as described in Appendix D.2 and thus, following (8), for a given transfer policy, the preferences of a municipality over two potential coalitions S and S' are a function of observed characteristics, parameters β , and idiosyncratic shocks.

Estimation will be via moment inequalities, as the solution to the coalition formation game presented in Section 1 is not unique, and thus a method that can handle nonuniqueness with an unknown selection rule is necessary. The moments used will ensure that, at the estimated parameters, the model neither predicts large amounts of merging where little actually occurred nor little merging where large amounts occurred. Identification is based on variation in observed covariates across municipalities and potential coalitions. In some areas, potential merger partners were geographically distant, while in other areas they were nearby. The transfer policy provided strong incentives for certain municipalities to merge in certain configurations, while having a much more limited effect in other cases. Population and related covariates also led to more substantial potential cost savings for some mergers than others.

The remainder of this section has the following form: first, the covariance structure of ϵ is described; then the details of each of the types of moment inequalities used are discussed; finally, the test statistic used is presented.

ber of parameters is challenging, because of both the computational costs involved in optimization and the tendency for confidence sets to become large in higher dimensions.

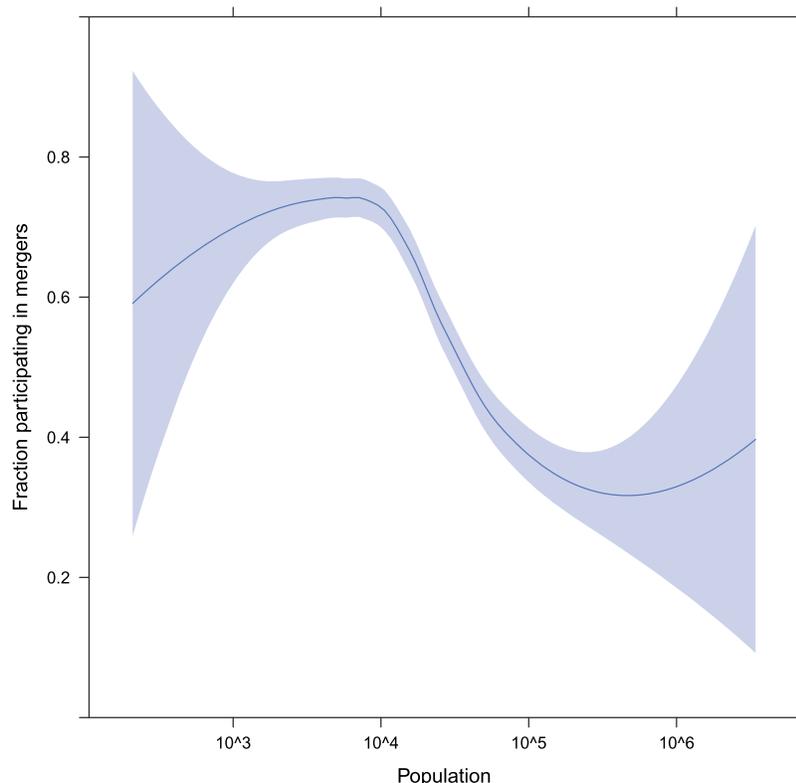


FIGURE 14. Mergers by population.

3.1 Structure of idiosyncratic shocks

It is implausible that ϵ is independent and identically distributed (i.i.d.) across different coalitions: if $S = \{m_1, m_2, \dots, m_{14}, m_{15}\}$ and $S' = \{m_1, m_2, \dots, m_{14}\}$, then a reasonable econometric model should have ϵ_S correlated with $\epsilon_{S'}$. The following construction will produce shocks such that $\epsilon_S \sim N(0, 1)$, not independent but identically distributed. Let \mathcal{S} be the set of all possible coalitions. The length of ϵ is then $|\mathcal{S}|$, which grows exponentially in the number of municipalities; however, the construction is such that all these shocks will be determined by a pair of random variables with length that grows only linearly in the number of municipalities.

The basic assumption used comes from the literature on ethnic fragmentation: under certain conditions, heterogeneous jurisdictions produce bad results for all residents, not only those far from the median voter.³³ First, suppose that for each individual resident there is an i.i.d. draw, $\omega_i \sim N(0, 1)$, representing i 's cultural identity. For municipality m with population N_m , the sample mean and sample variance of these draws will

³³While Japan is not known for extreme ethnic or linguistic heterogeneity, one could imagine even minor cultural differences playing such a role. [Costa and Kahn \(2003\)](#) provide a discussion of some of these issues.

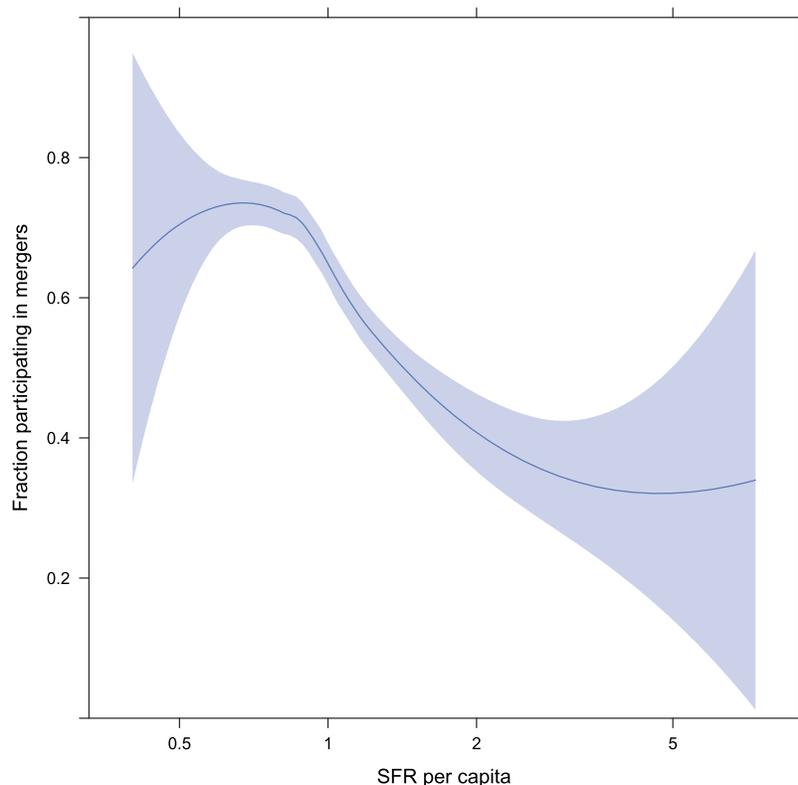


FIGURE 15. Mergers by revenue.

be

$$\bar{\omega}_m = \frac{1}{N_m} \sum_{i=1}^{N_m} \omega_i, \quad (18)$$

$$s_m^2 = \frac{1}{N_m - 1} \sum_{i=1}^{N_m} (\omega_i - \bar{\omega}_m)^2, \quad (19)$$

because there are N_m residents making i.i.d. draws. For a coalition S , define $\bar{\omega}_S$ and s_S^2 in the same way, replacing m with S . Define the vectors $\bar{\omega}_M$ and s_M to be the sample means and standard deviations for all municipalities. That is, $\bar{\omega}_M$ has length $|M|$, with an entry for each municipality m giving $\bar{\omega}_m$. Within-municipality heterogeneity is captured by the sample variance, so let

$$\epsilon_m = -f(X_m) \log s_m^2, \quad (20)$$

and likewise for any coalition S . Here $f(X) > 0$ is a function that generates weights such for any coalition $S = A \cup B$, then $f(X_S) > f(X_A)$ and $f(X_S) > f(X_B)$. That is, heterogeneity is relatively more important for larger municipalities. It is possible to calculate

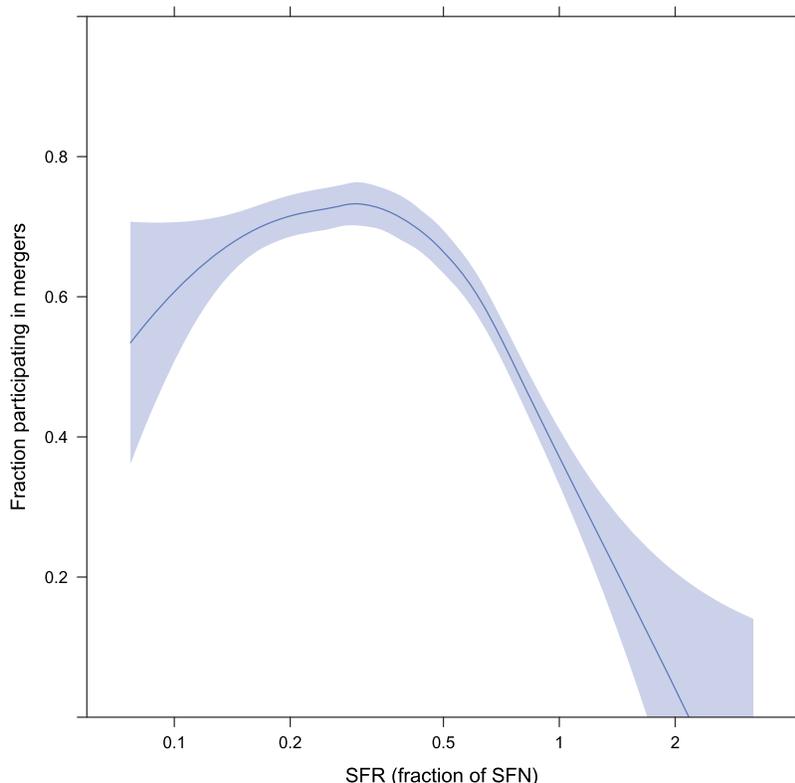


FIGURE 16. Mergers by fiscal strength.

ϵ_S for any coalition S given only $\bar{\omega}_M$ and s_M . Let $\epsilon(\bar{\omega}_M, s_M)$ be the vector of shocks resulting from such a calculation. If $f(X_S) = \sqrt{\frac{N_S-1}{2}}$, then $\epsilon_S \sim N(0, 1)$, not independent but identically distributed, as desired.³⁴ Further details are provided in Appendix E.1.

This particular structure for idiosyncratic shocks has not previously appeared in the literature. It is motivated by the difficulty of parameter estimation with a game that has multiple solutions and a choice set that grows exponentially in the number of players. To see why this setup is challenging, consider the case where a size 10 coalition is observed to form in a prefecture with 100 municipalities. There are $\binom{100}{10} \simeq 10^{13}$ potential such coalitions. If the shocks for these coalitions were i.i.d. normal and one size 10 coalition were observed in the prefecture, then the idiosyncratic term for that observed coalition could easily be 7 standard deviations into the positive tail of the distribution.

³⁴The idiosyncratic shock, taking the form of unobserved heterogeneity, resembles the (observed) geographic heterogeneity. There are two differences: the idiosyncratic shock is the same for all members of a coalition and the shock involves a scaling factor, f . The first of these is necessary to preserve computational feasibility; the second ensures that there always exists a set of idiosyncratic shocks that could rationalize any observed partition.

Assuming some degree of correlation between idiosyncratic shocks or imposing geographic restrictions on feasible coalitions reduces the magnitude of this problem but does not change its fundamental nature: for most actually observed partitions, every coalition in the partition could have a substantial positive idiosyncratic shock.³⁵ Moment inequality approaches of the type described in Pakes, Porter, Ho, and Ishii (2015) generally require idiosyncratic shocks that are mean zero conditional on the observed outcome, and thus these approaches do not extend easily to the model considered in this paper.³⁶

Ciliberto and Tamer (2009) style probability bounds can be applied to this problem and form an important part of the estimation strategy. In particular, it is relatively easy to compute a bound for the minimum amount of merging that should have been observed in a given area, regardless of the selection rule for choosing partitions from the stable set. However, it does not appear to be computationally feasible to compute a Ciliberto and Tamer style bound regarding the maximum amount of merging that should have been observed in a given area.³⁷ Parameter estimation using only this approach is thus difficult.

The form for the idiosyncratic shocks described above provides a solution to this problem. Specifically, the shocks have been constructed in such a way that the covariance matrix is effectively rank deficient. The rank of this matrix will grow linearly with the number of players in the coalition formation game, rather than exponentially as it would normally. This low rank covariance matrix makes it feasible to calculate a bound for the “least extreme” shocks that rationalize the observed partition.³⁸ These least extreme shocks can then be compared to the assumed distribution for the shocks. This

³⁵An additional problem with approaches of this sort is that it is not always clear which coalitions should be in \mathcal{S} . For example, no coalitions of size greater than 15 are observed in the data, but there was also no government policy that expressly prohibited a size 50 coalition from forming. Even a restrictive definition of potential mergers based both on number of participants and their geographic arrangement still leaves approximately 100 million possible coalitions, and ends up ruling out several coalitions that did occur. For example, there were about 10 mergers that were not geographically contiguous.

³⁶A “measurement error” approach—based on assuming that the difference between the outcomes predicted by the model and the actually observed outcomes is due to covariates measured with error—is also difficult to justify. The size of the outcome space vastly exceeds the number of covariates measured, and these outcomes do not appear to be ordered in such a way that measurement error alone could rationalize all potential partitions.

³⁷It is relatively easy to compute which coalitions have no refinements, but the obvious approaches for aggregating these coalition-level results into a bound regarding observed partitions appear to be computationally infeasible. There is also no obvious way to usefully classify partitions into a smaller number of “types” of outcomes and apply an approach such as Ciliberto and Tamer (2009) directly. Furthermore, unlike in Ciliberto and Tamer, where the same set of airlines play a game in various markets, it is not possible to match the players in a coalition formation game in one prefecture to those in another: in Kanagawa prefecture, the largest city and the second largest city are geographically adjacent, whereas in Shizuoka they are on opposite sides of the prefecture. Each prefecture has a new set of players with different characteristics, and the outcome of this particular coalition formation game is observed only once.

³⁸Calculating this bound requires considering the idiosyncratic shocks of all possible coalitions, which is computationally infeasible when the covariance matrix is full rank. The particular distribution assumed above, however, reduces the dimension of the shocks that need to be considered, making calculation of the bound feasible. A resulting weakness, unfortunately, is that the covariance structure of the error term cannot itself be estimated, as the model is only tractable for the specific structure assumed.

leads to moment inequalities that will penalize parameter values predicting merger patterns very different from those actually observed. In particular, moment inequalities of this type can be used to penalize parameter values that predict a maximum amount of merging that is less than the amount actually observed. An identified set can then be obtained by combining these moment inequalities with [Ciliberto and Tamer \(2009\)](#) style inequalities that penalize parameter values predicting a minimum amount of merging greater than the amount actually observed.³⁹

The moment inequalities related to the least extreme possible shocks are the methodological contribution of this paper, and thus are discussed first below. Following that discussion, three types of standard moments inequalities are presented: the model should not predict large numbers of mergers in the absence of the merger incentives, the model should not predict large numbers of mergers in prefectures where the actual merger incentives were quite weak, and the predicted tax rates in the model should not be wildly different from actually observed tax rates.⁴⁰

3.2 Moment inequalities

The first type of moment inequality used is based on the idea that at the true parameters β^0 , there should exist idiosyncratic shocks that both rationalize the observed coalition structure and are not too extreme compared to the distribution from which they were assumed to have been drawn. Begin by choosing an arbitrary function $h(\bar{\omega}, s|X)$. Define h^* as

$$h^*(\pi|X, \beta) = \min_{\bar{\omega}_M, s_M} h(\bar{\omega}_M, s_M|X) \quad \text{s.t.} \quad \pi \in \Pi^*(\epsilon(\bar{\omega}_M, s_M)|X, \beta), \quad (21)$$

where Π^* is the stable set from [Theorem 1](#) evaluated using the idiosyncratic shocks determined by $\bar{\omega}_M$ and s_M .

The restriction imposed in (21) is a stability restriction on partition π : specifically, h^* is a lower bound for h given that partition π was observed, and thus must have been part of the stable set Π^* . The idea here is that h is a function that returns, for a given $(\bar{\omega}, s)$ pair, a scalar value describing how extreme the shocks $\epsilon(\bar{\omega}, s)$ are, and h^* thus describes how extreme the least extreme shocks consistent with the observed partition are.⁴¹ It is

³⁹To use this approach with a different type of data, such as the networks described in [Jackson \(2008\)](#), it would be necessary to identify some parts of the data where the observed outcome is very well predicted by the structural terms of the model. Situations where there was a policy change with differential effects seem particularly likely to fulfill this condition. It is then necessary to show that if the structural term were not particularly important relative to the idiosyncratic term, the particular type of outcome observed would not have occurred with as high a frequency as it did. Showing this requires the ability to either simulate outcomes or calculate bounds, at least over the relevant part of the data. Situations where there are always a large number of possible solutions with very different characteristics would thus be difficult to estimate using the approach presented in this paper.

⁴⁰The idiosyncratic shocks are additively separable and do not depend on β . Specifically, the payoff for municipality m from merger S can be written as in (32): $u_{mS}(\beta) = v_{mS}(\beta) + \epsilon_S$.

⁴¹Good choices for h thus seem to be functions that give high values when $\bar{\omega}$ and s are extreme relative to the distributions from which they were assumed to have been drawn. The equations that follow are valid for any function h ; however, a poor choice for h will be ineffective in the sense that the resulting moment

always the case that

$$h(\bar{\omega}_M^0, s_M^0 | X) \geq h^*(\pi^0 | X, \beta^0), \quad (22)$$

where $\bar{\omega}_M^0$ and s_M^0 are the true values that were drawn for $\bar{\omega}_M$ and s_M , respectively, and π^0 is the partition that resulted from those draws. This is because $\Pi^*(\epsilon(\bar{\omega}_M^0, s_M^0) | \beta^0)$ must contain π^0 , otherwise π^0 could not have been observed: thus, $(\bar{\omega}_M^0, s_M^0)$ can be chosen in (21).

Now consider the moment

$$\begin{aligned} g_1(\pi, \beta | X) &= E_{\bar{\omega}_M, s_M} [h(\bar{\omega}_M, s_M | X)] - h^*(\pi | X, \beta) \\ &= \bar{h}(|X) - h^*(\pi | X, \beta). \end{aligned} \quad (23)$$

Here the $\bar{h}(|X)$ notation emphasizes that the expectation of $h(\bar{\omega}_M, s_M | X)$ can be calculated, as the distribution of $\bar{\omega}_M$ and s_M is known by assumption. The moment g_1 will be positive in expectation at the true parameter value β^0 , because

$$\begin{aligned} E_\pi [g_1(\pi, \beta^0 | X)] &= \bar{h}(|X) - E_\pi [h^*(\pi | X, \beta^0)] \\ &\geq \bar{h}(|X) - E_{\bar{\omega}_M, s_M} [h(\bar{\omega}_M, s_M | X)] = 0, \end{aligned} \quad (24)$$

where the second line follows from the first because inequality (22) holds at every $(\bar{\omega}_M, s_M, \pi)$ realization and thus also holds in expectation.

To see why this is the case, first note that π itself is random, even conditional on $(\bar{\omega}_M, s_M)$. This is because a given realization of $(\bar{\omega}_M, s_M)$ determines Π^* , but the rule that determines how the observed partition π is drawn from the stable set Π^* is unknown. The sequence of events is thus as follows. First, $\bar{\omega}_M$ and s_M are drawn, determining the idiosyncratic shocks. A partition π is then drawn from the resulting stable set $\Pi^*(\epsilon(\bar{\omega}_M, s_M) | X, \beta^0)$. As discussed above, for these draws of $\bar{\omega}_M, s_M$, and π , it must be the case that $h(\bar{\omega}_M, s_M) \geq h^*(\pi | X, \beta^0)$, because the actual $(\bar{\omega}_M, s_M, \pi)$ draw is guaranteed to satisfy the requirements of (21). For further clarity, a very simple example of this type of moment is given in Appendix E.5.

One problem with the g_1 moment is that it does not, in general, lead to a computationally tractable sample moment: h^* is a minimum over the idiosyncratic shocks corresponding to all the players present in the data being used to calculate the moment. For a specific choice of h , however, the calculation of h^* becomes additively separable across coalition formation games and can thus be calculated separately for each prefecture. Choose the form of h in (21) to be

$$\begin{aligned} h(\bar{\omega}_M, s_M | X) &= \sum_{m, m' \text{ adjacent}} \frac{N_m N_{m'}}{\sqrt{N_m + N_{m'}}} (\bar{\omega}_m - \bar{\omega}_{m'})^2 \\ &\quad + \sum_{m \in M} \frac{N_m - 1}{2} (s_m^2 - 1)^2, \end{aligned} \quad (25)$$

inequalities will lead to an identified set for $\hat{\beta}$ that is very large. Trivially, $h = 0$ will result in an identified set equal to the parameter space for β .

where the first summation is over pairs of municipalities that are geographically adjacent.⁴² The calculation of h^* in (21) is still not computationally feasible because of the very large number of potential deviations that need to be considered. However, as h^* is only used in the inequality in (22), a lower bound h^{**} can be used instead of calculating h^* directly. The calculation of this sort of bound is computationally feasible when the inequalities in (13) and (14) are used to check that a given partition is stable. Computational feasibility here is due to the specific construction of ϵ presented above, which allows any ϵ_S to be expressed as a linear function of a small number of random variables. Computational details for this type of moment inequality are provided in Appendix E.2.

The second type of moment inequality used is based on the following assumption: in the absence of any change in national government policy, merger activity in 1999–2010 should not have been greater than merger activity in 1979–1999. That is, assume that the increase in merger activity was caused by the change in national government policy. There is little debate in Japan that the large number of mergers that occurred during the 1999–2010 period were a result of policy changes made by the national government. Figure 1 shows that the merger activity is in marked contrast to the period before 1999: only 18 municipalities participated in mergers during the two decades preceding the implementation of merger promotion policies.

Let μ_Q^* be a lower bound on the number of mergers that the model predicts would have occurred in the absence of any government policy change.⁴³ Let Q be the number of municipalities involved in mergers in the 1979–1998 period. The moment

$$g_2(Q, \beta|X) = Q - \mu_Q^*(\beta|X) \quad (26)$$

can then be used as a moment inequality, with $E[g_2] \geq 0$ by the assumption just discussed. Computation, which is complicated by the need to account for the unknown selection rule, is explained in more detail in Appendix E.3.

The third type of moment inequality used is based on the right-hand side of Figures 8–11. Per capita government transfers are larger for smaller municipalities, resulting in the principal–agent problem, and thus the merger incentives offered, being of greater magnitude for those municipalities. Some prefectures, however, have mainly large municipalities and thus should not have been affected as strongly by the merger incentives; which exact prefectures these are will be discussed below.

Let Q^{99} be the number of municipalities actually participating in mergers in the 1999–2010 period, and let μ_Q^{*99} be a lower bound on the number of mergers that the model predicts would occur given the actual incentive policies in place.⁴⁴ The moment

⁴²This choice of h assigns very little weight to the sample means $\bar{\omega}$. The other moment inequalities will eliminate values of $\hat{\beta}$ where the model predicts many mergers that did not in fact occur. Thus, this moment inequality is used mainly to eliminate values of $\hat{\beta}$ that suggest that mergers that actually occurred should not have. As the $\bar{\omega}$ terms can only make sample variances higher, and thus in general will make mergers less attractive, the emphasis here is placed on the second term, the sample variances s^2 .

⁴³A lower bound must be used here rather than an exact prediction because the selection rule that chooses the observed partition π from the stable set Π^* is unknown.

⁴⁴Here too a lower bound needs to be used because the selection rule is unknown.

$g_3(Q, \beta|X) = Q - \mu_Q^*(\beta|X)$ can then be used as a moment inequality because

$$\begin{aligned} E_Q g_3(Q^{99}, \beta^0|X) &= \mu_Q^{99} - \mu_Q^{*99}(\beta^0|X) \\ &\geq 0 \end{aligned} \tag{27}$$

following the same argument as for the previous moment, with details as in Appendix E.3.

The fourth and final type of moment inequality used relies on the fact that actual tax rates are observed for all municipalities. This is particularly interesting in the merger period, where there is noticeable, although still low, dispersion in the tax rates being charged.⁴⁵ Specifically, suppose that the observed tax rates are a function of optimal tax rates plus some noise,

$$\tau_m^{**} = \max(\tau_m^*(\beta) + \varepsilon_m, \bar{\tau}), \tag{28}$$

where τ_m^* is taken from (5). If the theoretical model is correct, then including additional terms should not improve the fit of a Tobit regression. That is, if the restriction $\gamma = 0$ is imposed on the model

$$\tau_m^{**} = \max(\tau_m^*(\beta) + \gamma X_{mk} + \varepsilon_m, \bar{\tau}), \tag{29}$$

then if $g_4(\beta, X)$ is the gradient for γ , evaluated at $\gamma = 0$, this can be used as a moment equality.⁴⁶

3.3 Estimator

The four types of moments just presented could be used directly; however, there is substantial variation in observed characteristics of the (pre-merger) municipalities across prefectures. This is important because the large municipalities prevalent in Tokyo prefecture were not much affected by the merger policies, while the smaller (and poorer) municipalities in rural prefectures such as Tottori experienced a dramatic change in transfers. Sample moments that average across Tokyo and Tottori thus eliminate variation in observed municipal characteristics that would be helpful in obtaining precise estimates.

To preserve some of this variation in observables, divide prefectures into three groups: ones with fewer than 10% of municipalities having a population of less than 10,000 (metropolitan), ones with more than 65% of municipalities having a population of less than 10,000 (rural), and those in between. The summary statistics in Table 1 show that municipalities in metropolitan prefectures were much less likely to merge during the merger period than those in the other sorts of prefectures.

⁴⁵One complication here is that, de facto, municipalities appear not to be able to lower their tax rate below $\bar{\tau}$, although they are free to charge a higher rate. Even with this censoring, however, tax rates (after adjustment for the tax floor) should be correctly predicted by the model.

⁴⁶Here g_4 actually consists of two moment inequalities, because both an intercept and the size of municipalities in question is used with γ in (29), and both of these should have a coefficient of zero.

TABLE 1. Summary statistics (for municipalities, by type of prefecture).

	Units	Metropolitan	Mixed	Rural
\bar{c} (std. fiscal need)	¥100,000,000	361.86 (776.37)	64.42 (182.15)	53.85 (169.60)
$\bar{\tau}Y$ (std. fiscal rev.)	¥100,000,000	281.60 (665.12)	40.84 (146.98)	24.75 (112.47)
Population	100,000	2.04 (3.90)	0.34 (0.96)	0.23 (0.89)
Area (km ²)	100	0.45 (0.54)	0.98 (1.06)	1.93 (2.05)
City (vs. town/vill.)	%	70.83	19.51	16.00
MERGED_PRE_1999	%	1.39	1.41	0.00
MERGED_POST_1999	%	6.25	67.58	64.64
No. of municipalities		144	2486	625
No. of prefectures		3	36	8

Sample moments can then be computed for each of these groups. The third type of moment will only ever bind for metropolitan prefectures, due to the large number of mergers in the other types. The test statistic used for estimation is thus

$$\begin{aligned} \mathbf{T}(\beta) = & [\bar{g}_1^{\text{metro}}(\beta)]_-^2 + [\bar{g}_1^{\text{mixed}}(\beta)]_-^2 + [\bar{g}_1^{\text{rural}}(\beta)]_-^2 \\ & + [\bar{g}_2^{\text{metro}}(\beta)]_-^2 + [\bar{g}_2^{\text{mixed}}(\beta)]_-^2 + [\bar{g}_2^{\text{rural}}(\beta)]_-^2 \\ & + [\bar{g}_3^{\text{metro}}(\beta)]_-^2 + [\bar{g}_4(\beta)]_-^2, \end{aligned} \quad (30)$$

where \bar{g} are the sample moments, with $[x]_- = \min(x, 0)$.⁴⁷ The identified set and confidence set for $\hat{\beta}$ are then calculated for $\mathbf{T}(\beta)$ following [Andrews and Guggenberger \(2009\)](#), as described in Appendix E.4.

4. RESULTS

Results are shown in Table 2. The value of $\hat{\beta}_2$ gives the cost of geographic heterogeneity: an individual would be willing to have a municipal policy that was 1 km more distant in exchange for about ¥3500 per year. The value of $\hat{\beta}_3$ indicates that the national government labeling of \bar{c} as corresponding to a national minimum level of service appears to be correct: there is a fixed Stone–Geary style demand for 1.05 quality units of the public good, which is not statistically different from 1, but is statistically different from 0.

On the other hand, $\hat{\beta}_4$, the degree to which the central government's report \tilde{H} of efficiencies of scale in the provision of public services matches the true efficiencies of scale, is estimated to be about 0.5. That is, the national government reported efficiencies

⁴⁷The \bar{g}_4 sample moment is not broken down by prefecture types because it has relatively high variance; thus it is useful to use a large number of municipalities to calculate a single sample moment.

TABLE 2. Dependent variable is v_{mS} , (structural) utility to municipality m from merger S .

	Tax Floor	No Tax Floor
Consumption (β_0)	219.56** (145, 454)	118.18** (39, 348)
Government (β_1)	5.04** (2.6, 12.1)	2.80 (0.0, 9.8)
Distance (β_2)	-0.22** (-0.32, -0.04)	-0.23* (-0.39, 0.0)
STONE_GEAR (β_3)	1.04** (0.95, 1.12)	1.06** (0.97, 1.25)
EFF_OF_SCALE (β_4)	0.49** (0.38, 0.60)	0.51** (0.29, 0.66)
N (prefectures)	47	47

Note: **95% level. *90% level. (a, b) Extreme points for this variable in (five dimensional) 95% confidence set. Tax floor: municipalities cannot charge a tax rate less than $\bar{\tau}$, the national government's reference tax rate.

of scale that were roughly double the true efficiencies of scale.⁴⁸ Two null hypotheses can be rejected at very high confidence levels: that there are no efficiencies of scale in the provision of public goods ($\beta_4 = 0$) and that the efficiencies of scale in the provision of public goods are equal to the national government's \tilde{H}^0 report ($\beta_4 = 1$). As discussed in Section 2, a potential explanation for this exaggeration is that smaller municipalities are overrepresented in the national legislature.

Although the model is only set identified in theory, the minimizer of the test statistic is, in fact, a single point. This situation is standard in the literature, and at the estimated parameters, the model is not rejected at the 95% level.⁴⁹ The results thus show that an Alesina and Spolaore (1997) type model matches actually observed changes in political boundaries.

The theoretical model of Section 1 involves an objective function with curvature in transfers, and an associated unique first-best transfer scheme. The inefficiency discussed from a theoretical perspective in Cremer, De Kerchove, and Thisse (1985) is thus open to analysis from an empirical perspective using the parameter estimates just presented. In particular, how does the number of jurisdictions that will be observed differ from the optimal number and what are the welfare implications of this difference?

To answer this question, it is helpful to have a model of heterogeneity based on population distributed on a plane, as in Drèze, Le Breton, Savvateev, and Weber (2008). Specifically, an attractive feature of the geography-based heterogeneity used in Section 1 is that it predicts that there continues to be a trade-off between efficiencies of scale and

⁴⁸Estimates of efficiencies of scale in the provision of public goods are difficult to obtain (Reiter and Weichenrieder (1997)). This estimate is based on revealed preference over jurisdiction structure, and thus differs substantially from the traditional approach based on spending data from a fixed configuration of municipalities.

⁴⁹When a large number of moment inequalities are used, there may not be any parameter vector that satisfies all of them. This could be the result of idiosyncratic variation or could indicate that the model is misspecified.

heterogeneity, even as the area under consideration becomes large. For example, if the total population and surface area of Japan were increased by 1/47 by adding an additional prefecture, then the model would predict an increase in the optimal number of coalitions of about 1/47. In contrast, work following [Miceli \(1993\)](#), such as [Brasington \(2003\)](#), focuses on demographic characteristics. In these models, expanding the area under consideration would not increase the optimal number of jurisdictions proportionally unless the new area added had income levels and other characteristics different from any existing area.⁵⁰ Thus, the model used in this paper is particularly well suited for empirical consideration of the optimal number of jurisdictions.

Begin by considering the simplest possible case: population is uniformly distributed, all individuals are identical, there are no idiosyncratic shocks, and a social planner who weights everyone equally is drawing entirely new boundaries. An hexagonal arrangement of municipalities is optimal in this case, with all municipalities the same size.⁵¹ The planner's problem thus involves choosing a single size N_m for the municipalities so as to maximize the first three terms of (4),⁵²

$$N_m^* = \operatorname{argmax}_{N_m} \beta_0 \log\left(\frac{yN_m - \beta_3 c(N_m)}{yN_m}\right) + \beta_1 \log\left(\frac{yN_m - \beta_3 c(N_m)}{c(N_m)}\right) + \beta_2 \ell^{\text{hex}}(N_m), \quad (31)$$

where $\ell^{\text{hex}}(N_m)$ is the average distance of an individual to the centroid of his/her hexagon, given that the hexagon has population N_m and a population density the same as Japan (340 per km²).⁵³ This formula yields an optimal municipal population of 199,988. This suggests 635 municipalities for all of Japan, compared to the 1750 actually present at the end of the merger period.

The actual Japanese population distribution, however, is not uniform, and some areas may have been too remote to merge. Suppose that the national government were to impose centralized mergers, rather than allow municipalities to choose. Assume that the T^0 transfer policy will be used for the resulting amalgamated municipalities. Finding the optimal pattern of centralized mergers is a difficult combinatorial problem, but

⁵⁰This issue does not cause empirical problems in [Brasington \(2003\)](#), as only adjacent jurisdictions are considered for estimation, and legal restrictions on mergers justify this approach.

⁵¹The optimality of hexagonal tiling appears to have been known since at least the early 20th century: [Drèze et al. \(2008\)](#) provide extensive references. For populations on bounded areas there are some issues, as hexagons do not, in general, tile well near the boundary. In the case considered here, however, the municipalities that would be on the perimeter are a relatively small fraction of total municipalities, and thus this problem is ignored.

⁵²The assumption here is that the government will not provide any transfers because all municipalities will have the same cost function (it can be shown that this indeed is the case at the optimum). Spending on the public and private good thus follows the standard expenditure share pattern for a Stone–Geary utility function.

⁵³The actual cost function used for computation is $c(N_m \bar{x})$, where \bar{x} is the average per capita level of various characteristics. That is, if 20% of the population of Japan is over 65 years old, then 0.20 N_m people in each municipality will be assumed to be over 65 years old for the cost calculation.

a relatively simple “greedy” algorithm yields a rough approximation.⁵⁴ To perform this calculation, though, it is necessary to know the weights w that the national government places on various municipalities. Consider three potential sets of weights.

For the first potential set of weights, assume that, as discussed in Section 2, the grey values in Figure 6 correspond to the true cost function c , but the national government also included additional adjustments so as to create the black reported costs \tilde{c}^0 . As \tilde{c}^0 was in use during the period in which there were basically no mergers, assume that it is the solution to the fixed-boundary optimum transfers from (10). For each municipality, an implied weight \hat{w}_m can then be calculated by comparing \tilde{c}^0 to the assumed c .

As a second option, suppose that the weights w happen to be such that the national government wished to misreport c in exactly the way given in (16). The implied weights \hat{w} can then be calculated by using $\hat{\beta}_4$ and (16) directly. A third option is that \tilde{c} was chosen for arbitrary and inexplicable reasons, and the national government weights all municipalities equally in per capita terms.⁵⁵

Using the greedy approximation for the optimal centralized mergers, 993 municipalities are optimal for the first set of weights, 1097 for the second set, and 796 for the third set.⁵⁶ Thus, for plausible weights, the optimal centralized merger pattern is close to the announced target of 1000 municipalities and is lower than the actual outcome of 1750 municipalities.

To calculate the inefficiency resulting from the decentralized mergers, begin by considering the mergers that were actually observed and suppose that the transfer policy had remained unchanged at T^0 .⁵⁷ For now, ignore the idiosyncratic term. Depending on the weights used, the benefit to the national government from the mergers that were actually observed is between ¥222 billion and ¥432 billion.⁵⁸ If the national government had instead carried out centralized mergers as discussed above, the benefits would range from ¥520 billion to ¥912 billion. Thus, setting aside the idiosyncratic term, the actual mergers provided only 43–47% of the gains that would have occurred if the government had instead implemented centralized mergers. This percentage may appear low given that the number of municipalities did fall from 3255 to 1750; however, the decentralized mergers resulted in some bizarre arrangements. The most notorious of these is a set of six geographically contiguous municipalities in Aomori prefecture that

⁵⁴Begin with a partition of all singleton coalitions. Find the coalitions S and S' such that replacing them with $S'' = S \cup S'$ would produce the greatest benefit. Repeat, now with the partition including S'' , until there are no more mergers that will improve the objective function. Each iteration reduces the number of coalitions by 1, and thus the algorithm is guaranteed to terminate (either with the grand coalition or some finer partition).

⁵⁵The equal per capita weight is calculated as the (population weighted) average of the weights obtained using the first option.

⁵⁶If $\beta_3 = 1$ and $\beta_4 = 1$ are used instead of the estimated values for $\hat{\beta}_3$ and $\hat{\beta}_4$, then the optimal number of municipalities is only 387.

⁵⁷Unfortunately it appears difficult to apply an equation such as (31) to determine the number of mergers in a decentralized case: for example, Haimanko, Le Breton, and Weber (2007) show that even in a relatively simple setup, the number of jurisdictions does not change monotonically with the polarization of the population distribution.

⁵⁸Ignoring the idiosyncratic term, the gains to the national government from the actual mergers with the actual T^1 transfer policy range from ¥34 billion to ¥292 billion for the weights considered above.

merged into three noncontiguous pairs.⁵⁹ Thus, many of the mergers that actually occurred were not the most attractive from the national government's perspective, at least based on observed characteristics.

Calculating the inefficiency of mergers with respect to the actual data in a way that includes the idiosyncratic error term is challenging because the actual coalition formation game is not computationally feasible to simulate, and even if it were feasible, the selection rule for choosing a partition from the stable set Π^* is unknown. Some idea of the importance of the idiosyncratic shock can be obtained by considering the case where the national government carried out centralized mergers with perfect information. For the first set of weights, the gains from centralized mergers with perfect information would be ¥1097 billion, whereas centralized mergers without knowledge of ϵ result in gains of only ¥650 billion. Excluding the idiosyncratic term, the actually observed mergers yielded about half the gains that would have been obtained from centralized mergers.⁶⁰ If it were also true that, including the idiosyncratic term, the actually observed mergers yielded about half the gains that would have been obtained from (perfect information) centralized mergers, then the decentralized mergers that actually occurred resulted in gains about equal to (asymmetric information) centralized mergers. In this case, from the national government's perspective, decentralized mergers were as good a choice as centralized mergers, but both resulted in substantial inefficiency relative to a hypothetical set of perfect information centralized mergers.

A remaining question is why there were relatively few mergers observed in the decentralized case. One possibility is that the asymmetric information problem is severe and the optimal merger incentives are quite weak. Another possibility is that the merger incentives actually offered were weak relative to the optimal incentives. To answer this question, a description is needed of the optimal merger incentives for the national government to offer in the context of the model in Section 1. The potential for multiple stable partitions, along with the unknown selection rule, makes it difficult to compute the optimal transfer scheme for the national government to offer in the coalition formation game that actually occurred; however, below a solution is provided in the two-municipality case. In this case the optimal incentives turn out to be much stronger than the incentives actually offered.

5. OPTIMAL MERGER INCENTIVES

5.1 Theory

Consider the model in Section 1, where municipalities have the option of merging and the national government does not observe ϵ . This setup is that of a principal–agent problem, with the agents' utility entering directly into the principal's objective rather than indirectly through a participation constraint. The agents choose an action (merge or not merge), and the principal observes the action but not some of the information (ϵ) that

⁵⁹The anecdotal explanation here is that adjacent municipalities differed in per capita tax base, while the mergers that occurred involved participants that were similar on this dimension.

⁶⁰The gains from centralized mergers with imperfect information are between 44% and 62% of the gains from centralized mergers with perfect information, depending on weights.

led the agents to make that choice. Depending on the action taken, the agents' benefit from transfers may be higher (no merger) or lower (merger). The principal wants to set a transfer policy that makes higher per capita transfers in the case where they are more helpful, but doing so results in the agents sometimes no longer choosing the action the principal would prefer. The result will be that the principal chooses a second-best set of transfers so as to lessen this distortion.

Focus on the case with two municipalities, m and m' , and a cost function that exhibits efficiencies of scale. The national government is assumed to commit to a transfer policy: first, the national government chooses $T = \{T_m, T_{m'}, T_S\}$, and the municipalities then choose whether to merge or not, with the merger S occurring if both municipalities are in favor. One option is for the national government to set T_m and $T_{m'}$ to be extremely negative, resulting in the municipalities choosing to merge regardless of ϵ . A better outcome might be possible, however, with higher values for T_m and $T_{m'}$, as the municipalities will then choose to merge when $\epsilon_S - \epsilon_m$ and $\epsilon_S - \epsilon_{m'}$ are high, and not merge otherwise.⁶¹

Let v be all the structural terms in the utility function in (7):

$$u_{mS}(T_S) = v_{mS}(T_S) + \epsilon_S \tag{32}$$

and similarly for u_{mm} .

THEOREM 2. *The optimal transfer policy T^* is (implicitly) described by*

$$\frac{(T_m^{FB} - T_m^*) + (T_{m'}^{FB} - T_{m'}^*)}{T_S^* - T_S^{FB}} = \frac{\Pr(\{S\}|T^*)}{1 - \Pr(\{S\}|T^*)}, \tag{33}$$

$$\frac{(T_m^{FB} + T_{m'}^{FB} - T_S^{FB}) + \eta \frac{w_{m'}}{b} e' + (1 - \eta) \frac{w_m}{b} e}{(T_m^{FB} - T_m^*) + (T_{m'}^{FB} - T_{m'}^*) + (T_S^* - T_S^{FB})} = 1 + \frac{\Pr(\{S\}|T^*)(1 - \Pr(\{S\}|T^*))}{(\beta_0 + \beta_1) \frac{\partial \Pr(\{S\}|T)}{\partial v_S} \Big|_{T^*}}, \tag{34}$$

$$\frac{\frac{T_{m'}^{FB} - T_{m'}}{\partial \Pr(\{S\}|T^*)} - \frac{T_m^{FB} - T_m}{\partial \Pr(\{S\}|T^*)}}{\frac{\partial v_{m'}}{\frac{w_{m'}}{b} e' - \frac{w_m}{b} e}} = \frac{\beta_0 + \beta_1}{1 - \Pr(\{S\}|T^*)}, \tag{35}$$

where

$$\eta = \frac{\frac{\partial \Pr(\{S\}|T)}{\partial v_m}}{\frac{\partial \Pr(\{S\}|T)}{\partial v_S}} \tag{36}$$

⁶¹Even better outcomes would be possible in the most general case where, following the revelation principle, the national government would offer a transfer schedule (and merger decision) based on the municipalities' reports of ϵ . The optimal schedule is not obvious, and difficulty in determining it may be one reason why incentives of this sort do not appear to be used in connection with municipal mergers. The more restricted optimal case examined below provides a better match with actual government policy.

is the fraction of marginal mergers that occur because of a change in the merger decision of municipality m , and

$$e = E[u_{mS} - u_{mm} | u_{m'S} = u_{m'm'}, u_{mS} \geq u_{mm}, T] \quad (37)$$

is the expected benefit to municipality m from a marginal merger for municipality m' , with e' being the reverse.

See Appendix E.1 for the proof.

Here the three incentive terms $(T_S^* - T_S^{\text{FB}})$, $(T_m^{\text{FB}} - T_m^*)$, and $(T_{m'}^{\text{FB}} - T_{m'}^*)$ will all be positive. The national government benefits from municipal mergers more than the municipalities themselves, and thus will offer additional transfers beyond the fixed boundary T_S^{FB} when the municipalities merge. Conversely, it will choose a policy that is not as generous as the fixed boundary T_m^{FB} and $T_{m'}^{\text{FB}}$ for the case where the municipalities do not merge.

The transfers here are a standard carrot and stick policy: the national government wants to provide an incentive for municipalities to merge and it has two types of transfer amounts it can vary. The local government utility function is curved, and thus the national government uses both the carrot (setting T_S^* to be greater than T_S^{FB}) and the stick (setting T_m^* and $T_{m'}^*$ to be less than T_m^{FB} and $T_{m'}^{\text{FB}}$).

Equation (33) shows that the degree to which the stick is used relative to the carrot is proportional to the fraction of municipalities merging. This is because, conditional on a given merger decision, inefficiency arises from the gap between the fixed-boundary optimal transfer and the chosen policy only for the transfer that actually occurs: T_S^* if the municipalities merged versus T_m^* and $T_{m'}^*$ if they did not. If municipalities almost always merge, then the stick can be threatened but rarely used; conversely, if municipalities almost never merge, then the carrot can be offered but rarely given. The transfer that is more often actually paid out thus should be closer to the fixed-boundary optimum, with the other transfer providing most of the merger incentive.

Equation (34) describes the overall magnitude of the incentive. The left-hand side is the benefit of a merger relative to the merger incentive being offered. The $\frac{\Pr(\{S\}|T)(1-\Pr(\{S\}|T))}{\partial \Pr(\{S\}|T)/\partial v_S}$ term on the right-hand side is related to the “thinness” of the distribution of ϵ at the merger/no merger threshold. Intuitively, the carrot and stick are only useful when they change municipalities’ merger decisions. As the distribution of ϵ becomes more and more spread out, the incentive has less of an effect on municipal merger decisions, but the inefficiency arising from the difference between the actual transfer and the fixed-boundary optimum remains. Thus, the national government will offer weaker incentives as the idiosyncratic shock becomes more important in municipal decision-making.

Equation (35) describes how the stick portion of the incentive is distributed across the two municipalities. The right-hand side is the relative strength of the stick incentive: as described for (33), this is higher when municipalities are more likely to merge. On the

left-hand side, the numerator is the incentive offered relative to the rate at which municipalities are convinced to merge. The denominator is the benefit of a marginal merger to the other municipality. The intuition here is that the national government wants to offer an incentive to a municipality when it will change that municipality's decision and result in a merger, and when this new merger will benefit the other municipality.

Simple closed form solutions for T^* do not appear to be available for any standard probability distribution. Numerical calculation, however, is possible:

EXAMPLE 2. There are two identical municipalities, each of which is a rectangle with length:width ratio of $1 : \sqrt{2}$ and uniform population.⁶² The idiosyncratic shocks are $\epsilon_m = \epsilon_{m'} = 0$ and $\epsilon_S \sim \text{Logistic}(0, \sigma)$. The national government weights individuals in both municipalities equally. The cost function is $c(N_m) = c_0 + c_1 N_m$.

With these idiosyncratic shocks $\frac{\Pr(\{S\}|T)(1-\Pr(\{S\}|T))}{\partial \Pr(\{S\}|T)/\partial v_S} = \sigma$, making it easy to solve numerically for the optimal policy for a given set of parameters. Set parameters to plausible values and consider optimal national government transfer policies for pairs of municipalities with initial population N_m .⁶³

Figure 8 shows the shape of the cost function in this setup, graphed in per capita terms. Per capita cost declines more quickly in population at low populations, and thus $2T_m^{\text{FB}} - T_S^{\text{FB}}$ will be larger per capita in this region.⁶⁴ The difference between the national government benefit from a merger and the local government benefit is thus largest in this region, with the result that the national government will offer stronger merger incentives for these smaller municipalities. Figures 10 and 12 show the optimal national government policy, expressed as a fraction of the cost function: this is $\frac{T_m^{\text{FB}} - T_m^*}{c(N_m)}$ for the stick and $\frac{T_S^* - T_S^{\text{FB}}}{c(N_m)}$ for the carrot. Thus, in Figure 10, a value near zero shows that the national government is choosing T_m^* to be very close to the fixed-boundary optimum T_m^{FB} , while a value of -0.4 indicates that the government is punishing municipalities that do not merge by reducing their transfer by an amount that they could use to produce 0.4 units of the public good.⁶⁵

⁶²The amalgamated municipality will be geometrically similar to the initial municipalities: it is a rectangle with a length:width ratio of $\sqrt{2} : 1$ and uniform population density.

⁶³For parameters β_0 , β_1 , and β_2 , use values from the first column of Table 2, which will be discussed in Section 4. Set $\beta_3 = 1$ and $\beta_4 = 1$. For the cost function, use $c_0 = \text{¥}1.3$ billion and $c_1 = \text{¥}130,000$, which are roughly the ordinary least squares (OLS) coefficients from a regression on unadjusted standard fiscal need data (discussed in Section 2). Use population density equal to the average for Japan, use y equal to GDP per capita, and choose b/w such that the optimal government transfer is equal to the cost of providing the Stone–Geary minimal level of quality.

⁶⁴As the municipalities are identical, restrict consideration to national government policies that have $T_m = T_{m'}$. With identical municipalities and a single idiosyncratic shock, an alternative form for the left-hand side of (34) is $(1 + \frac{T_S^* - 2T_m^*}{2T_m^{\text{FB}} - T_S^{\text{FB}}})^{-1}$. This form makes it clear that as the idiosyncratic shock becomes very small, the optimal policy becomes $T_S^* = 2T_m^*$, where, following (33), $T_S^* = T_S^{\text{FB}}$ if the merger will occur and $T_m^* = T_m^{\text{FB}}$ if it will not.

⁶⁵The leftmost point (close to -0.47) of the line in Figure 10 corresponds to the lowest population for which the policy given by (33) and (34) is actually a solution. At this point, T_m^* is such that τ_m^* is close to 1 and any further reduction in T_m thus forces a merger.

The merger incentive for the smallest municipalities is provided mostly via the stick, as these municipalities merge at very high rates. On the other hand, larger municipalities merge at lower rates, and the incentive for them thus emphasizes the carrot. Very large municipalities would experience little change in efficiencies of scale after a merger, meaning that the fixed-boundary optimal transfers would be roughly the same in both the merger and no merger cases. For very large municipalities, then, national government and local government objectives are roughly aligned, and thus the merger incentives the national government offers are very small.

5.2 Comparison with actual policy

In a principal–agent setup, asymmetric information results in the optimal incentives being such that municipalities do not always merge when the national government would prefer them to do so. Specifically, large idiosyncratic shocks lead to optimal incentives being relatively weak, as shown in (34). This would then result in the pattern of decentralized mergers being less extreme than what the national government would implement if it could observe ϵ .

To consider this empirically, continue to assume that the old T^0 transfers were the fixed-boundary optimum transfers. An index \mathcal{I}_S of the strength of the merger incentives that were actually offered for a merger $S = \{m, m'\}$ can then be obtained by considering the reciprocal of the left-hand side of (34):

$$\mathcal{I}_S = \frac{(T_m^0 - T_m^1) + (T_{m'}^0 - T_{m'}^1) + (T_S^1 - T_S^0)}{T_m^0 + T_{m'}^0 - T_S^0}. \quad (38)$$

Here the horizontal externality terms e and e' are assumed to be zero, resulting in \mathcal{I} being higher than would be the case otherwise. If the true e and e' were included in the denominator, then \mathcal{I} should be in $[0, 1]$, with lower values of \mathcal{I} corresponding to the optimal incentives for a larger idiosyncratic shock.

Consider the merger $S = \{m, m'\}$, where m has a smaller population than m' . The average incentive \mathcal{I} offered for a merger involving a pair of municipalities with populations N_m and $N_{m'}$ is shown in Figure 17.⁶⁶ Figure 18 shows the distribution of \mathcal{I} for the mergers that actually occurred.⁶⁷ As the strength of the optimal incentive is decreasing in the importance of the idiosyncratic shock, the actual incentive policy may be optimal if idiosyncratic shocks are quite important in determining which municipalities merge.

To evaluate this possibility, consider an environment consisting of all adjacent pairs of municipalities, where each of these 7438 pairs are treated independently. With $\hat{\beta}$ and a choice of implied weights \hat{w} , the optimal government transfers T^* characterized by (33)–(35) for any given pair can be calculated numerically.

⁶⁶The incentive is indicated by the color at the point corresponding to N_m on the horizontal axis and $N_{m'}$ on the vertical axis. Most points are white or close to white, indicating an incentive of about 0.5.

⁶⁷Equation (38) is extended appropriately for mergers of size greater than 2.

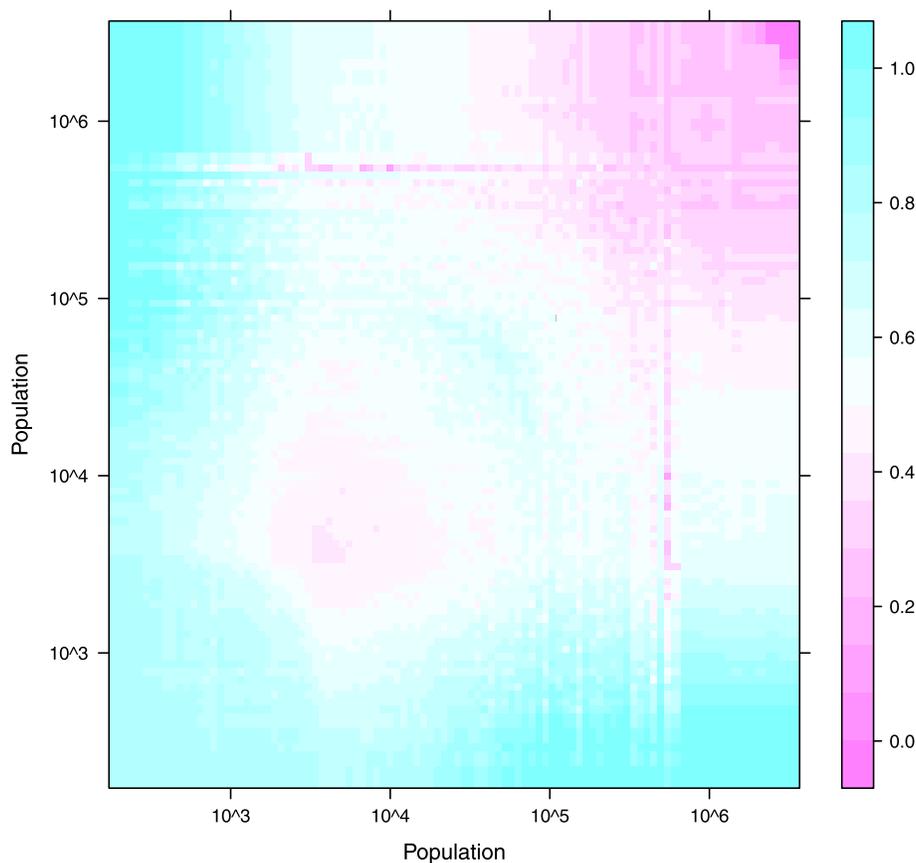


FIGURE 17. Actual incentive (random pairs). Values displayed are an average over randomly selected pairs of municipalities in the relevant population bins. The high values on the extreme left-hand side and the low values on the right-hand side correspond to populations where there are few actual municipalities.

Using the first set of weights described above, at T^{FB} in expectation, 4% of the pairs would choose to merge. If the national government could observe ϵ , however, it would want 85% of the pairs to merge. If the national government could not observe ϵ , but had to choose centrally which mergers would occur, then it would merge 88% of the pairs.⁶⁸ With imperfect information, the optimal T^* chosen by the national government is such that 83% of the pairs merge. Thus, in the pairs setup, there does not appear to be that substantial an effect of imperfect information on the number of mergers that occur.

Calculating the incentive index I from (38) for T^* with the pairs data set gives some idea as to why this might be the case. The optimal incentives actually exceed 1 on this index, because of the presence of horizontal externalities: it will usually be the case that one municipality benefits more from a merger than the other. Even in the case

⁶⁸This is larger than the full information optimum because for most pairs, merging is in expectation a good choice from the perspective of the national government.

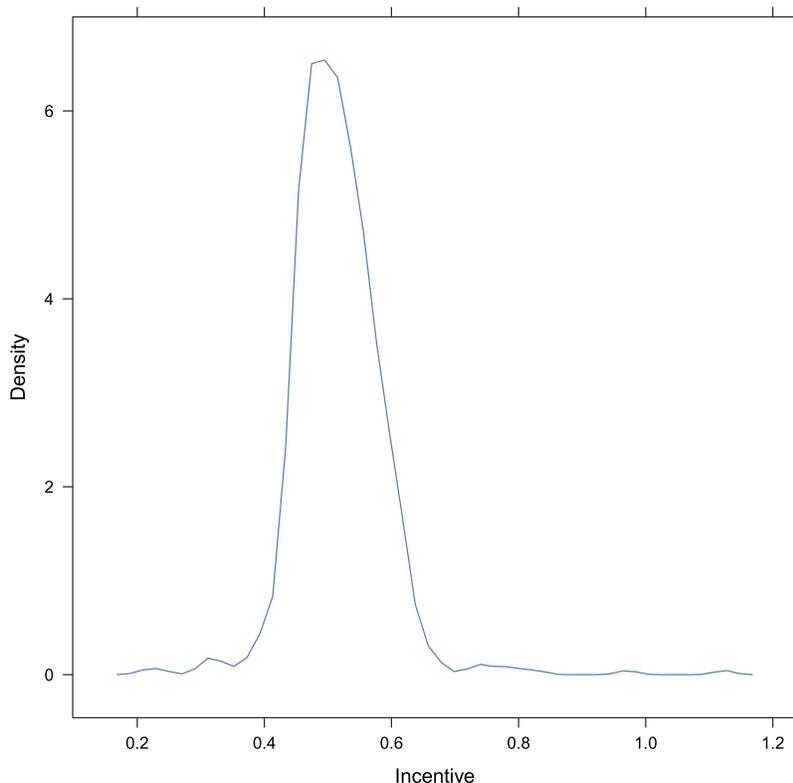


FIGURE 18. Actual incentive (actual mergers).

where these externalities are eliminated, though, by considering identical municipalities with perfectly correlated idiosyncratic shocks, the median optimal incentive is 0.91. The optimal policy thus provides a merger incentive equivalent to most of the gains from the merger, and roughly the same number of mergers occur as in the centralized case.

An incentive of about 0.9 is substantially higher than the 0.53 average for the actual incentives. For the optimal incentive to be about 0.5, there would not only have to be no horizontal externalities, but the idiosyncratic shock would have to have a standard deviation about 7 times higher relative to the structural parameters than is the case for the estimated $\hat{\beta}$. With this large an idiosyncratic shock, the difference between the T^{FB} policy and the T^* policy is no longer particularly important, because mergers occur basically at random.

It thus appears that the actual incentives offered by the Japanese national government were weak relative to the optimal incentives. To test whether this could explain the missed merger target, first compute the index \mathcal{I} for each pair of municipalities using the actual transfer policy. Then consider a merger incentive that had the optimal form, but only the magnitude given by this \mathcal{I} . In this case, between 22% and 52% of pairs would choose to merge, compared to between 66% and 83% for an incentive with the optimal

magnitude. The gains from this weak incentive are only 27% to 51% of the gains from an optimal incentive.

It thus appears that the gap between the stated 1000 municipality objective and the actual 1750 municipality merger pattern could be explained by merger incentives that had only the observed magnitude, which is low relative to the optimum. A remaining question is what might have led the national government to offer incentives that were weak relative to the optimal incentives. Appendix F.2 considers potential explanations for this difference. It appears that it is not driven by restrictions on the transfer function the national government can use. However, calculations show that the optimal incentives for smaller municipalities involve either transfer payments that differ from the fixed-boundary optimum transfers by 25% forever or a lump sum payment that amounts to 75% of GDP per capita. Implementing incentives this large may not have been feasible if there were some limit on the national government's ability to commit to a transfer policy.

6. CONCLUSION

This paper examined a set of decentralized municipal mergers in Japan and found that the post-merger number of municipalities was higher than the national government would have chosen if it had implemented a centralized merger policy instead. The low number of mergers that occurred appears to be due to the merger incentives offered by the national government, which are weaker than the optimal incentives.

These conclusions rely on parameter estimates obtained using a structural model of local governments and an estimator based on moment inequalities. Estimation was possible due to a special covariance structure assumed for the error term and the use of a new moment inequality based on the variance of the idiosyncratic shock. The structural model suggested a relevant index by which to measure the strength of merger incentives, and simulation using the estimated parameters was then used to compare the optimal incentives to those actually in place.

APPENDIX A: STABILITY CONCEPT

Suppose that player $m \in M$ has preferences \preceq_m defined over the set $\{S \subset M \mid m \in S\}$, with \prec_m indicating a strict preference. Extend these preferences to partitions in the following way: if $\pi(m)$ is the coalition that municipality m belongs to in partition π , then $\pi \preceq_m \pi'$ if $\pi(m) \preceq_m \pi'(m)$. Let $\pi \prec_S \pi'$ for some coalition S if $\forall m \in S, \pi \prec_m \pi'$.

The theoretical literature dealing with a coalition formation problem of this form is extensive: for example, the consistent set of [Chwe \(1994\)](#) and the farsighted stable set of [Diamantoudi and Xue \(2003\)](#) are both applicable. However, as noted by [Osmani \(2011\)](#), the computation required for finding solutions in a [Chwe \(1994\)](#) style setup is exponential in the number of players, and it appears a similar result would hold for the farsighted stable set. To maintain computational feasibility even with large numbers of players, then, the solution set is defined using the [Von Neumann and Morgenstern \(1944\)](#) (VNM) stable set.

Although the VNM stable set was originally defined in terms of imputations rather than coalition structures, this paper follows Ehlers (2007) and others in defining the stable set over coalition structures. Specifically, the Von Neumann–Morgenstern solution requires that (i) no coalition structure in the stable set be dominated by another coalition structure in the set and that (ii) any coalition structure outside of the set is dominated by a coalition structure belonging to the set.

DEFINITION 1 (Ehlers (2007)). Let $<$ be a dominance operator and let $\Pi^{\text{VNM}} \subseteq \Pi$. Then Π^{VNM} is called a stable set for $(\Pi, <)$ if the following two properties hold:

1. For all $\pi, \pi' \in \Pi^{\text{VNM}}$, $\pi \not< \pi'$ (internal stability).
2. For all $\pi \notin \Pi^{\text{VNM}}$, $\exists \pi' \in \Pi^{\text{VNM}}$ where $\pi < \pi'$ (external stability).

Ray and Vohra (1997) only allow deviating coalitions to force refinements of a partition, and Diamantoudi and Xue (2007) show that this creates a stable set.⁶⁹ The hedonic game considered in this paper is simpler than the “equilibrium coalition structures” that Ray and Vohra examine; thus in this paper, both deviations involving refinements (splits of coalitions) and coarsenings (mergers of coalitions) will be allowed. Otherwise, the theory follows that presented in Ray and Vohra. Let $\pi \nearrow_S \pi'$ and $\pi \searrow_S \pi'$ mean that $\pi <_S \pi'$, $S \in \pi'$, where π' is a coarsening and a refinement of π , respectively. Using the terminology of Ray and Vohra, π is blocked by π' if either there is a set of coalitions in π that are unanimously in favor of merging to create π' or there is a subset of “perpetrators” in π that are unanimously in favor of deviating from their current coalition. In the former case, π' is the coarsening that results from the merger, while in the latter case it is a refinement that includes a coalition for these perpetrators and some arrangement of the “residual” left behind when the perpetrators deviated, such that the configuration of perpetrators and residual is stable. More formally, where \rightarrow should be read as “blocked by,”

DEFINITION 2. We have $\pi \rightarrow \pi'$ if $\exists S$ such that either $\pi \nearrow_S \pi'$ or $\pi \searrow_S \pi'$, where these are defined as follows:

1. We have $\pi \nearrow_S \pi'$ if $\pi' \setminus \pi = S$ such that $\pi <_S \pi'$, and
 - (a) $S = \bigcup Q$ for some $Q \subset \pi$,
 - (b) $\nexists S' \subset S$ such that $\pi' \searrow_{S'} \pi''$.
2. We have $\pi \searrow_S \pi'$ if $\exists S \in \pi'$ such that $\pi <_S \pi'$, and
 - (a) $\pi \setminus \pi' = S'$ with $S' = \bigcup Q'$ for some $Q' \subset \pi'$,
 - (b) $\nexists \tilde{Q}$ such that $Q' \rightarrow \tilde{Q}$.

⁶⁹An alternative approach would be to allow only single player deviations, as in Greenberg (1979). Ray and Vohra (1997) is used instead because anecdotal evidence suggests that multiplayer deviations involving a refinement or a coarsening were more common than single player deviations not to a refinement or a coarsening during the coalition formation process.

The recursion is well defined since Q' is a proper subset of π' . Now let \rightarrow be the transitive closure of \rightarrow .⁷⁰ Assume that $\Pi \neq \emptyset$.

PROPOSITION 1. *We have $\Pi^* = \{\pi \mid \nexists \pi' \text{ such that } \pi \rightarrow \pi'\}$ is a stable set with respect to (Π, \rightarrow) .*

PROOF. By construction, Π^* is internally stable. Now take some $\pi \notin \Pi^*$. Then $\exists\{\pi_1, \dots, \pi_n\} \subset \Pi$ such that $\pi \rightarrow \pi_1 \rightarrow \dots \rightarrow \pi_n$ and either $\pi_n \in \Pi^*$ or there is a cycle with $\pi_n = \pi_l$ for some $l < n$. If there is such a cycle, then it must contain both mergers and dissolutions. However, such a cycle cannot exist because \nearrow is defined such that there are no refinements. \square

The proof of Theorem 1 in the main text is then very straightforward.⁷¹

PROOF OF THEOREM 1. Existence. By the above definition of Π^* , existence is immediate.

Nonemptiness. If $\Pi \setminus \Pi^* = \emptyset$, then Π^* is not empty because Π is assumed not to be empty. If $\Pi \setminus \Pi^* \neq \emptyset$, then Π^* is not empty because external stability was shown in the proof of Proposition 1.

Uniqueness. Suppose that Π^{**} is also a stable set with respect to (Π, \rightarrow) . Consider the bipartite directed graph defined by \rightarrow with $\Pi^{**} \setminus \Pi^*$ and $\Pi^* \setminus \Pi^{**}$ as the two sets of nodes. Every node must have in-degree of at least 1, but there can be no cycles. The only such graph is empty and, thus, $\Pi^{**} = \Pi^*$.⁷² \square

APPENDIX B: COST ESTIMATES

The official government formula for the calculation of $\tilde{c}(X_m)$ is

$$\tilde{c}(X_m) = \sum_{k=1}^{24} X_{mk} \cdot \tilde{c}_k (1 + \tilde{H}_k(X_m)). \quad (39)$$

Here the public good is viewed as a sum of 24 component goods, such as fire fighting, care for the elderly, resident registration, and so forth.⁷³

⁷⁰That is, $\pi \rightarrow \pi'$ if either $\pi \rightarrow \pi'$ or $\exists\{\pi_1, \dots, \pi_n\}$, where $\pi \rightarrow \pi_1 \rightarrow \dots \rightarrow \pi_n \rightarrow \pi'$. To see why the transitive closure is used here, consider the case where $\pi_1 \searrow_S \pi_2 \nearrow_{S'} \pi_3$. The π_1 and π_2 should not be in the stable set, while π_3 should, but $\{\pi_3\}$ is not a VNM stable set with respect to \rightarrow because $\pi_1 \rightarrow \pi_3$.

⁷¹It can also be shown that Π^* contains a Pareto optimal partition. All partitions in Π^* , including those that are not Pareto optimal, are treated equally, since imposing additional restrictions at this stage would mean that the solution set would no longer be the outcome of the cooperative game coalition formation process described above. There may be some solutions that seem particularly unattractive: $\{\pi \in \Pi^* \mid \exists \pi' \in \Pi^*, \pi \rightsquigarrow \pi'\}$. While the theory above could likely be rewritten to shrink the stable set, eliminating these elements, it would be computationally infeasible to use any of these new restrictions in the empirical section, as they would require enumerating the entire stable set.

⁷²To see this, attempt to iteratively construct a nonempty graph that has the desired form.

⁷³Fire fighting is “metropolitan government” responsibility in the special ward (i.e., old Tokyo City) area of Tokyo prefecture. Special wards did not participate in municipal mergers, however, and thus the discussion that follows does not cover the *sui generis* government in place there.

Each of these component goods is associated with a target population X_{mk} , an estimated unit cost \bar{c}_k , and an adjustment coefficient \tilde{H}_k . The target population for services to the elderly is the number of residents over 65, the target population for agricultural services is the number of farmers, and so forth. The unit cost is calculated based on the cost of providing component good k to one member of the target population for k when m is the “reference municipality”: a hypothetical city with a population of 100,000, surface area of 160 km², and other standard characteristics. The adjustment coefficient is created by multiplying and adding together a set of (usually) decreasing splines determined by X_m .⁷⁴ These adjustments generally result in higher per capita cost estimates for smaller municipalities.

One pattern frequently observed is that \tilde{H}_k takes the form

$$\tilde{H}_k(X_m) = \prod_{j \in J_1} \tilde{H}_k^j(X_{mj}) + \frac{1}{X_{mk} \bar{c}_k} \sum_{j \in J_2} \tilde{H}_k^j(X_{mj}). \quad (40)$$

The total number of available adjustment coefficients available in $J_1 \cup J_2$ is 15, but all 15 are never used for the same component good k .⁷⁵ One interesting feature here is that the adjustments based on characteristics in J_2 do not actually depend on the unit cost that they are supposedly adjusting, due to the division by $X_{mk} \bar{c}_k$.⁷⁶ Thus, de facto, the method for calculating $\tilde{c}(X_m)$ is

$$\tilde{c}(X_m) = \sum_{k=1}^{24} X_{mk} \cdot \bar{c}_k (1 + \tilde{H}_k(X_m)) + \zeta_m, \quad (41)$$

where

$$\tilde{H}_k(X_m) = \prod_{j \in J_1} \tilde{H}_k^j(X_m). \quad (42)$$

Of the adjustment coefficients in J_1 , by far the most important is the *dankai* (literally step or grade) adjustment, which is based on the scale of the service provision. The *dankai* adjustment is generally based on the target population for the service in question, which for most services is the total number of residents. This adjustment is substantial, with the per capita cost of providing services usually estimated to be 2–3 times higher for a municipality of 4000 people than one of 100,000. *Dankai* adjustments for some important services are shown in Figure 19.

There are also other adjustments, such as one for population density: for example, the estimate for the cost of providing fire fighting is increased from ¥1.009 million to

⁷⁴The X_m is more than a 24-tuple, with some elements used only in the calculation of the adjustment coefficient.

⁷⁵The precise number of adjustment coefficients and component goods varies slightly from year to year.

⁷⁶Expenses in this second group include those related to different sorts of land (forest, farmland, etc.) and costs related to American military bases. More problematically, they also include subsidies related to construction bonds issued earlier in the 1990s. The values of ζ used in this paper attempt to ignore the subsidies on these old outstanding bonds, but as exact data on bond payment schedules are not available, it is difficult to do this perfectly.

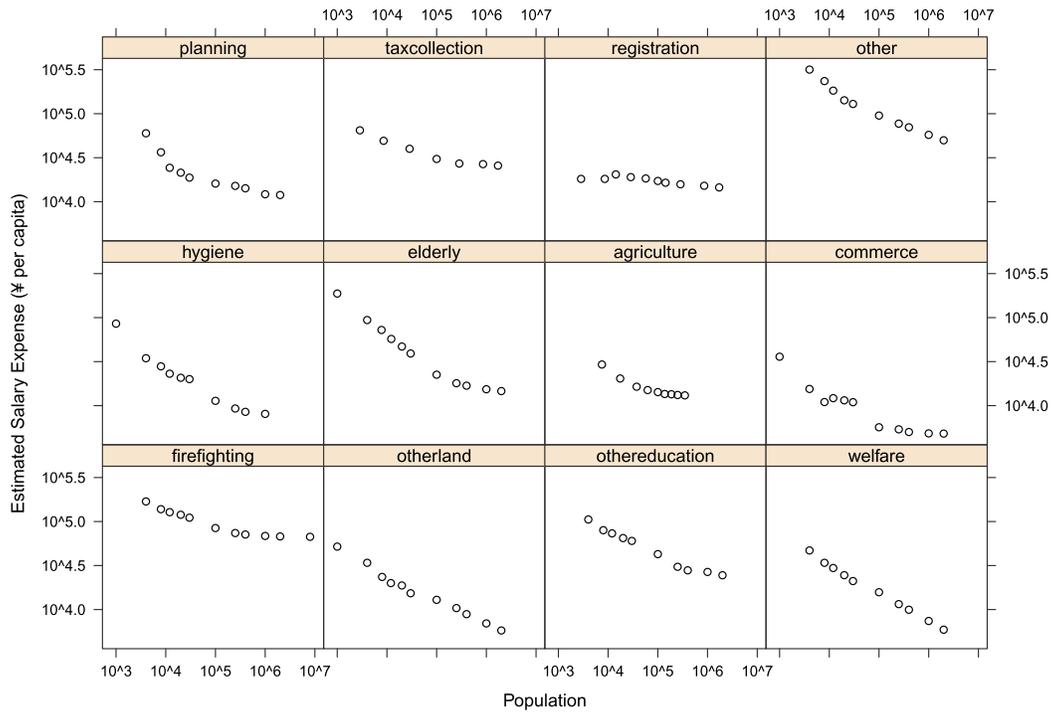


FIGURE 19. Adjustment coefficient based on number of personnel required. (Vertical axis is coefficient multiplied by estimated “standard” personnel expenditures.)

¥1.029 million if the population density of a city with population 100,000 were 150 per km² rather than 200. Population density also affects the estimated cost of other services such as elderly care and resident registration, but in different ways, with the effect on fire-fighting costs in general being greater than the effect on other component goods.⁷⁷

To match this system to the theoretical model presented in Section 1, assume that each municipality must choose a single quality q at which to provide all the component public goods: it is not possible, for example, for municipality m to choose to provide quality $q_m^f = 5$ fire fighting, but only quality $q_m^r = 3$ resident registration. Furthermore, suppose that the cost estimate $\tilde{c}(X)$ produced by the national government has the right general form, but with possibly incorrect adjustments \tilde{H} . Specifically, suppose that the true cost of one quality unit of the public good is

$$c(X_m) = \sum_{k=1}^{24} X_{mk} \cdot \tilde{c}_k \cdot (1 + H_k(X_m)) + \zeta_m \tag{43}$$

⁷⁷It was not possible to obtain data regarding some of the smaller adjustment coefficients. For the analysis conducted in this paper this portion of the costs are included in ζ_m . Thus, the data used in this paper actually slightly understate the Ministry’s estimates of efficiencies of scale, making the finding that $\beta_4 < 1$ more surprising.

and, as outlined in (16), the national government reports a possibly different cost

$$\tilde{c}(X_m) = \sum_{k=1}^{24} X_{mk} \cdot \bar{c}_k \cdot (1 + \tilde{H}_k(X_m)) + \zeta_m, \quad (44)$$

where $\tilde{H}_k = \psi H_k$.

Municipal financial information was obtained from the Ministry of Internal Affairs and Communications. The unit costs \bar{c}_k and adjustment coefficients \tilde{H}_k were more challenging to obtain, both due to the complexity of the formulae and the fact that some of the data used in the calculations are not publicly available. Discussions with Ministry officials confirmed that formulae for \tilde{H}_k are determined by the expert opinion of Ministry officers, and are not created directly via a regression of municipal characteristics on previous municipal spending or by applying a specific set of a priori assumptions regarding efficiencies of scale.⁷⁸

The process by which Ministry officials actually calculate \tilde{c} involves first calculating the number and type of local bureaucrats necessary to provide each of the component services. The cost of equipment and materials, plus any transfers to the relevant target population (e.g., child benefit payments), is then added. Finally, the number and type of bureaucrats that smaller and larger municipalities would require to provide the same quality of service is calculated.⁷⁹ National Personnel Authority salary scales are used to convert employee numbers to a total wage bill, which is added to an adjusted estimate for equipment and materials. By definition there are no economies of scale with respect to transfers to individuals, since the same level of service would imply the same level of transfers in the cases where there are transfer payments.

Ministry calculations of \bar{c}_k and \tilde{H}_k are subject to two types of outside interference. First, the amount of transfers allocated needs to somehow match the budget agreed upon with the Finance Ministry. This appears to be accomplished by modifying capital spending estimates, with the result that official municipal capital spending “needs” vary radically from year to year; estimates of the noncapital spending required to provide municipal services, on the other hand, change very little.⁸⁰ This sort of variation can be captured in the model presented in Section 1 via a change in b , the cost of public funds.

A second sort of interference comes from politicians, as well as line ministries such as the former Construction Ministry, and involves pressure to promote spending on local projects. Over time, this resulted in the addition of numerous “project” adjustment

⁷⁸According to Ministry of Internal Affairs and Communications officials, each year estimates are modified based on formal and informal feedback from municipalities and prefectures, observed spending patterns, and in-house research.

⁷⁹The sizes at which these estimates are performed varies slightly from year to year and from service to service, but in recent years estimates have generally been produced for populations of 4000, 8000, 12,000, 20,000, and 30,000 for municipalities below the reference size, and at 250,000, 400,000, 1,000,000, and 2,000,000 for municipalities above the reference size.

⁸⁰Occasionally modifications are also made by adding additional expense categories. These are distinguishable from the usual expense categories by their placement at the end of the list of expenses, their short lifespan, and their nonspecific names. The usual expense categories have remained effectively unchanged since at least 1968.

coefficients, each providing a special incentive for a variety of public works projects. DeWit (2002) describes the history of this interference, which makes it clear that government estimates of capital spending requirements are not closely related to actual costs. This conclusion is supported by actual capital spending patterns, which are not at all close to government estimates. This sort of variation is captured in the model presented in Section 1 through a β_4 that is less than 1, indicating that the government is exaggerating expenses. The idea that a local government might be forced by the national government to spend money on public services that it does not want is captured through a tax floor at $\bar{\tau}$, one of the specifications estimated in Section 4.

B.1 Merger policies

Incentives for municipalities to merge were provided in the form of both a stick and a carrot. The stick policy involved a substantial reduction in transfers, focused particularly on smaller municipalities. The national government substituted \tilde{H}^1 for \tilde{H}^0 , resulting in a \tilde{c}^1 that was substantially lower than \tilde{c}^0 for low population municipalities. Figure 9 shows the difference between \tilde{c}^0 and \tilde{c}^1 that resulted from this change, expressed as a fraction of \tilde{c}^0 .

This reduction was mainly due to two policy changes.⁸¹ First, a “flattening” was applied to adjustment coefficients for municipalities with a population of less than 4000. The announced objective was that \tilde{H} should not increase above the level calculated for a municipality of 4000 people.⁸² In reality, however, this was applied only to some component services, and the components that were not affected by this policy were some of those that had the greatest curvature in \tilde{H}_k . Thus, although according to announced policy, the shape of the \tilde{c} curve shown in Figure 7 should have changed substantially, with the left-hand side becoming flat, in reality the flattening policy had a less dramatic effect. A second policy involved slight reductions in adjustment coefficients for municipalities between 4000 and 100,000 population. The combined effect of these policies is quite similar to a change in ψ of the type shown in Figure 5.

The national government’s stated reason for this decrease in transfers was to remedy budgetary deficits. An interesting feature of the change, however, is that there was a relatively limited decrease for larger municipalities: less than 5% for almost all municipalities over 250,000 population. A decrease this small is not consistent with a substantial increase in the cost of public funds, b . Thus, although officially the change from \tilde{c}^0 to \tilde{c}^1 was unrelated to municipal mergers, the lack of cuts to larger municipalities combined with the similarity of Figure 9 to Figure 10 suggests that the change to \tilde{H} was de facto the stick half of the national government’s merger incentives. The effect of the shift to \tilde{c}^1 was widely reported in the Japanese popular press, often in the form of complaints that the smallest municipalities were being forced to merge.⁸³

⁸¹These changes took place around 2003 but were announced earlier.

⁸²For component services k that had a target population X_{mk} different from the total population, the announced objective was that the adjustment coefficient \tilde{H}_k should not increase beyond the level associated with a reference target population calculated for a municipality of total population 4000.

⁸³This is not literally true: in the actual data, even if $c = \tilde{c}$ and $\beta_3 = 1$, there is always some tax rate at which a municipality could have provided at least quality β_3 of public services and thus could remain

An interesting feature of the stick policy is that, although the left tail of the S-curve in Figure 10 does not show up clearly in Figure 9, this flattening out is apparent in the changes in underlying splines from \tilde{c}^0 to \tilde{c}^1 . The flattening is not visible in the data because there are so few municipalities with populations that are sufficiently small. Furthermore, the few municipalities that are in this region are remote islands or other “special” areas, and thus have a substantial amount of noise in the financial statistics from various special programs.

The carrot part of the merger incentives was provided by two policies. The first of these allowed municipalities to keep more of the savings from a merger. Specifically, starting in 1999 a “substitute calculation” was offered that used $\sum_{m \in S} T_m$ rather than T_S to calculate transfer payments. This is equivalent to a benefit of

$$\sum_{m \in S} T_m - T_S \quad (45)$$

and was provided for 10 years starting from the date of the merger.⁸⁴ This benefit is exactly equal to the reduction in transfers that would have otherwise occurred as a result of the merger. Thus, whereas previously municipalities would normally see a decrease in transfers following a merger, this would no longer be the case, at least for the first decade following the merger.

The second carrot policy was the *Gappei Tokureisai*—special subsidized bond issues allowed for amalgamated municipalities. These bonds were to be issued in the 10 years following the merger, and the total amount allowed was a function of the number of municipalities participating in a merger and their populations. The value of these bonds is calculated based on the subsidy offered, using information from Ishihara (2000) and prefectural government sources. The official explanation for the special merger bonds was to eliminate any direct financial cost of merging, such as the construction of a new city hall; however, the bonds appeared to allow significant capital expenditures beyond the actual costs of amalgamation. The calculation used in this paper assumes zero transition costs related to building new facilities. If such costs actually existed, this would lead to a higher transfer being optimal for merging municipalities. The conclusion of this paper is that incentives were not high enough, and the assumption that there were no such costs is thus conservative.

Figure 11 shows the hypothetical case where the carrot policies continued forever and each municipality merged with an identical copy of itself.⁸⁵ In reality, however, the carrot policies just discussed disappeared 10 years after the merger. To fit this policy to the static model presented in Section 1, municipalities will be assumed to smooth

independent. The meaning of the media reports appears to be that mergers that previously appeared very unattractive had suddenly become attractive, because the tax rate and quality of service in the singleton case had worsened.

⁸⁴Some limited merger incentives were offered even prior to the merger period. In particular, the benefit described in (45) was provided for 5 years after a merger. This was extended to 10 years during the merger period. In both periods there was also an intermediate amount provided for a “phase-out period” of 5 years.

⁸⁵The figure shows $T_S^1 - T_S^0$, expressed in units of \tilde{c}^0 .

consumption via borrowing and (more relevantly) saving, with the optimal smoothed spending amount determined via a present value calculation.⁸⁶

The special merger policies just described began to be phased out in 2006, which motivated many municipalities to finalize mergers in 2005.⁸⁷ By this point there were only 1844 municipalities remaining, as shown in Figure 1, down from 3255 at the start of the merger period. A small number of mergers occurred during the phase-out period, reducing the final number of municipalities to 1750 in 2010. For estimation purposes, these mergers are treated as though they were finalized prior to 2006 and implementation was simply delayed for exogenous reasons.⁸⁸

During the merger period, a separate set of policies known as the *Trinity Reforms* were also implemented, shifting some taxes and associated expenditures from the national government to local governments. These changes were accompanied by a modest reduction in unit cost \bar{c}_k for many component goods k : this is the source of the reduction in \bar{c} for large municipalities, visible in Figure 9. This reduction does not appear to have a substantial effect as a merger incentive, but is included in the calculation of \tilde{c}^1 . The shift in tax base does not affect merger incentives at all, as it was counterbalanced by a decrease in transfers, and municipalities charged exactly the standard rate on the transferred tax base. If, in the future, municipalities began charging a different rate on this transferred tax base, then there would be a very small effect from the change in revenues. This effect is not considered in this paper.

REFERENCES

- Acemoglu, D. (2003), “Why not a political Coase theorem? Social conflict, commitment, and politics.” *Journal of Comparative Economics*, 31 (4), 620–652. [259]
- Alesina, A., R. Baqir, and W. Easterly (1999), “Public goods and ethnic divisions.” *Quarterly Journal of Economics*, 114 (4), 1243–1284. [260]
- Alesina, A., R. Baqir, and C. Hoxby (2004), “Political jurisdictions in heterogeneous communities.” *Journal of Political Economy*, 112 (2), 348–396. [260]
- Alesina, A. and E. Spolaore (1997), “On the number and size of nations.” *Quarterly Journal of Economics*, 112 (4), 1027–1056. [257, 260, 286]
- Alesina, A. and E. Spolaore (2003), *The Size of Nations*. MIT Press, Cambridge, MA. [260]
- Andrews, D. W. K. and P. Guggenberger (2009), “Validity of subsampling and ‘plug-in asymptotic’ inference for parameters defined by moment inequalities.” *Econometric Theory*, 25 (3), 669–709. [285]

⁸⁶The interest rate used for these calculations is 2%, the rate at which municipalities generally issued bonds.

⁸⁷The situation in Shizuoka prefecture at that time is shown in Figure 4.

⁸⁸Explaining why a coalition would not form during the 1999–2005 period, but would under the progressively less-advantageous policies in place in 2006–2010 would require adding elements to the model, such as arrival of new information, that would substantially complicate the analysis. This paper thus treats the entire 12 years as a single period.

- Baccara, M., A. İmrohoroğlu, A. J. Wilson, and L. Yariv (2012), “A field study on matching with network externalities.” *American Economic Review*, 102 (5), 1773–1804. [259]
- Banerjee, S., H. Konishi, and T. Sönmez (2001), “Core in a simple coalition formation game.” *Social Choice and Welfare*, 18 (1), 135–153. [262]
- Barberà, S. and A. Gerber (2007), “A note on the impossibility of a satisfactory concept of stability for coalition formation games.” *Economics Letters*, 95 (1), 85–90. [265]
- Bogomolnaia, A. and M. O. Jackson (2002), “The stability of hedonic coalition structures.” *Games and Economic Behavior*, 38 (2), 201–230. [259]
- Bogomolnaia, A., M. Le Breton, A. Savvateev, and S. Weber (2008), “Stability of jurisdiction structures under the equal share and median rules.” *Economic Theory*, 34 (3), 525–543. [265]
- Brasington, D. M. (2003), “Snobbery, racism, or mutual distaste: What promotes and hinders cooperation in local public-good provision?” *Review of Economics and Statistics*, 85 (4), 874–883. [287]
- Chwe, M. S.-Y. (1994), “Farsighted coalitional stability.” *Journal of Economic Theory*, 63 (2), 299–325. [296]
- Ciliberto, F. and E. Tamer (2009), “Market structure and multiple equilibria in airline markets.” *Econometrica*, 77 (6), 1791–1828. [280, 281]
- Costa, D. L. and M. E. Kahn (2003), “Civic engagement and community heterogeneity: An economist’s perspective.” *Perspective on Politics*, 1 (1), 103–111. [277]
- Cremer, H., A.-M. De Kerchove, and J.-F. Thisse (1985), “An economic theory of public facilities in space.” *Mathematical Social Sciences*, 9 (3), 249–262. [286]
- De Donder, P., M. Le Breton, and E. Peluso (2012), “On the (sequential) majority choice of public good size and location.” *Social Choice and Welfare*, 39 (2–3), 457–489. [260]
- Demange, G. (1994), “Intermediate preferences and stable coalition structures.” *Journal of Mathematical Economics*, 23 (1), 45–58. [260]
- Desmet, K., M. Le Breton, I. Ortuño-Ortín, and S. Weber (2011), “The stability and breakup of nations: A quantitative analysis.” *Journal of Economic Growth*, 16 (3), 183–213. [260]
- DeWit, A. (2002), “Dry rot: The corruption of general subsidies in Japan.” *Journal of the Asia Pacific Economy*, 7 (3), 355–378. [302]
- Diamantoudi, E. and L. Xue (2003), “Farsighted stability in hedonic games.” *Social Choice and Welfare*, 21 (1), 39–61. [296]
- Diamantoudi, E. and L. Xue (2007), “Coalitions, agreements and efficiency.” *Journal of Economic Theory*, 136 (1), 105–125. [297]
- Diermeier, D., H. Eraslan, and A. Merlo (2003), “A structural model of government formation.” *Econometrica*, 71 (1), 27–70. [265]

- Dreze, J. H. and J. Greenberg (1980), "Hedonic coalitions: Optimality and stability." *Econometrica*, 48 (4), 987–1003. [263]
- Drèze, J., M. Le Breton, A. Savvateev, and S. Weber (2008), "'Almost' subsidy-free spatial pricing in a multi-dimensional setting." *Journal of Economic Theory*, 143 (1), 275–291. [286, 287]
- Ehlers, L. (2007), "Von Neumann–Morgenstern stable sets in matching problems." *Journal of Economic Theory*, 134 (1), 537–547. [297]
- Farrell, J. and S. Scotchmer (1988), "Partnerships." *Quarterly Journal of Economics*, 103 (2), 279–297. [263]
- Fox, J. T. (2010), "Identification in matching games." *Quantitative Economics*, 1 (2), 203–254. [259]
- Gale, D. and L. S. Shapley (1962), "College admissions and the stability of marriage." *American Mathematical Monthly*, 69 (1), 9–15. [265]
- Gordon, N. and B. Knight (2009), "A spatial merger estimator with an application to school district consolidation." *Journal of Public Economics*, 93 (5–6), 752–765. [260, 263, 265]
- Greenberg, J. (1979), "Stability when mobility is restricted by the existing coalition structure." *Journal of Economic Theory*, 21 (2), 213–221. [297]
- Greenberg, J. and S. Weber (1986), "Strong Tiebout equilibrium under restricted preferences domain." *Journal of Economic Theory*, 38 (1), 101–117. [260]
- Haimanko, O., M. Le Breton, and S. Weber (2004), "Voluntary formation of communities for the provision of public projects." *Journal of Economic Theory*, 115 (1), 1–34. [260]
- Haimanko, O., M. Le Breton, and S. Weber (2005), "Transfers in a polarized country: Bridging the gap between efficiency and stability." *Journal of Public Economics*, 89 (7), 1277–1303. [264]
- Haimanko, O., M. Le Breton, and S. Weber (2007), "The stability threshold and two facets of polarization." *Economic Theory*, 30 (3), 415–430. [288]
- Hirota, H. (2007), "Verification of municipal incorporation with discrete choice model." *Planning Administration*, 30 (4), 75–81. [In Japanese.] [273]
- Hirota, H. (2012), "Heisei no daigappei ni okeru shichouson no sentaku gyoudou no ken-shou." In *Chihou Bunken he no Chousen: Atarashii Koukyou no Keizai Bunseki* (S. Saito, ed.), 99–120, Osaka University Press, Suita. [273]
- Horiuchi, Y. and J. Saito (2003), "Reapportionment and redistribution: Consequences of electoral reform in Japan." *American Journal of Political Science*, 47 (4), 669–682. [268]
- Ishihara, N. (2000), *Shin Chihou Zaisei Chousei Seido Ron*, new edition. Gyousei, Tokyo. [303]

Jackson, M. O. (2008), *Social and Economic Networks*. Princeton University Press, Princeton, NJ. [281]

Kido, H. and E. Nakamura (2008), “Strategy and environment: Determinants of municipal amalgamations.” *The Annals of the Japanese Society for Public Administration*, 43, 112–130. [In Japanese.] [273]

Miceli, T. J. (1993), “The decision to regionalize in the provision of education: An application of the Tiebout model.” *Journal of Urban Economics*, 33 (3), 344–360. [287]

Osmani, D. (2011), “A note on computational aspects of farsighted coalitional stability.” Working Paper FNU-176, Sustainability and Global Change Research Unit, Hamburg University. [296]

Pakes, A. (2010), “Alternative models for moment inequalities.” *Econometrica*, 78 (6), 1783–1822. [259]

Pakes, A., J. Porter, K. Ho, and J. Ishii (2015), “Moment inequalities and their application.” *Econometrica*, 83 (1), 315–334. [280]

Ray, D. and R. Vohra (1997), “Equilibrium binding agreements.” *Journal of Economic Theory*, 73 (1), 30–78. [265, 297]

Reiter, M. and A. Weichenrieder (1997), “Are public goods public? A critical survey of the demand estimates for local public services.” *FinanzArchiv: Public Finance Analysis*, 54 (3), 374–408. [286]

Roth, A. (2008), “Deferred acceptance algorithms: History, theory, practice, and open questions.” *International Journal of Game Theory*, 36 (3), 537–569. [259]

Von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ. [296]