ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208269

New Haven, CT 06520-8269

http://www.econ.yale.edu/~egcenter

Economic Growth Center Discussion Paper No. 1022

Economics Department Working Paper No. 113

Political Mergers as Coalition Formation: An Analysis of the *Heisei* Municipal Amalgamations

Eric Weese

Department of Economics

Yale University

April 2013

Notes: Center discussion papers are preliminary materials circulated to stimulate discussion and critical comments.

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, Tal Gross, Masayoshi Hayashi, Hidehiko Ichimura, Vadim Marmer, Konrad Menzel, Masashi Nishikawa, Tai Otsu, Nancy Qian, Pablo Querubin, 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, based on initial computations performed at MIT. The usual disclaimer applies.

Political Mergers as Coalition Formation*

An Analysis of the *Heisei* Municipal Amalgamations

Eric Weese†

April 2013

Abstract

Due to moral hazard problems, municipal mergers in Japan did not result in as many amalgamations as a central planner would have chosen. The inefficiency of the decentralized mergers is calculated using structural parameter estimates based on observed mergers and actual national government policies. Estimation requires neither an equilibrium selection assumption nor the enumeration of all possible mergers.

JEL codes: C63, D71, H77

Keywords: Boundaries, Mergers, Moment Inequalities, Municipalities

Self-determination is intuitively appealing. However, 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

*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, Tal Gross, Masayoshi Hayashi, Hidehiko Ichimura, Vadim Marmer, Konrad Menzel, Masashi Nishikawa, Tai Otsu, Nancy Qian, Pablo Querubin, 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, based on initial computations performed at MIT. The usual disclaimer applies.

†Economics Department, Yale University. Email: eric.weese@yale.edu.

boundaries.[1] Currently, there is limited 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 more than twice as large as it would have been under centralized decision-making.

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.[2] 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. Fiscal tightening at the end of the 1990s led to the national government announcing a special merger-promotion transfer policy. The nature of this policy 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 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 tradeoff 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

[1] Despite two thirds support for independence in a referendum (99% yes with about a third of the population boycotting), the Republic of South Ossetia is currently recognized by only four countries. After a referendum with even higher support and a 17 year delay, the Republic of Kosovo is recognized by about half of UN members. At the subnational level, there is variation in the treatment of municipal boundary changes even between very similar countries. For example, in Finland and Norway municipal mergers were voluntary, in Sweden they were forced, and in Denmark there appears to have been a hybrid system.

[2] For comparisons, ¥1=1¢ is a rough but useful approximation. During the period in which financial data is analyzed, the USD/JPY exchange rate has varied from ¥147=$1 (Aug. '98) to ¥80=$1 (Oct. '10). GDP per capita has remained relatively constant at ¥4,000,000.

costs and benefits of potential configurations. The possibility of municipal mergers thus leads to the potential for inefficiency, as mergers that are optimal from the local perspective may not be from the national perspective.

The merger process is modelled 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.[3] 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.

These problems are dealt with by using an estimator based on moment inequalities, following Pakes [2010]. Necessary conditions for the observed mergers to be stable can be checked without enumerating all other potential mergers, and these conditions must hold regardless of whether the observed outcome is only one of many stable partitions. Most of the moment inequalities used in estimation are of the type employed in Ciliberto and Tamer [2009] and Pakes, Porter, Ho, and Ishii [2011], but one moment inequality based on the variance of the idiosyncratic shocks has not previously appeared in the literature. A specific choice of covariance structure, based on interpreting the idiosyncratic term as unobserved cultural heterogeneity, reduces the dimension of the error term and makes estimation feasible. A resulting weakness of the model is that the covariance structure of the error term cannot itself be estimated, as the model is only tractable for the specific structure assumed.

Parameter estimates show a tradeoff between geographic proximity and efficiencies of scale in the provision of local public services. The estimated efficiencies of scale, however, are not as extreme as the estimates used by the national government in its calculation of transfer payments. If this paper's estimates of efficiencies of scale are correct, then the transfer formula actually used by the national government is inconsistent with it weighting all individuals equally; however, this transfer formula is consistent with the national government maximizing an objective function where indi-

[3] 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 of Fox [2008] and other recent papers analyzing coalition formation games relevant to industrial organization.

vidual weights are proportional to a simple measure of legislative malapportionment at the national level.

Using this additional result regarding the national government's objective function, numeric estimates for inefficiency from the perspective of the national government can be calculated. First, in order to provide an incentive for municipalities to engage in decentralized mergers, the national government had to offer a special transfer policy different from its first-best policy for fixed boundaries: the inefficiency here is equivalent to ¥250 billion in national government spending. The mergers that resulted from this incentive scheme, however, are equivalent to an additional ¥335 billion in national government spending even with the second-best transfers. It is also possible to compare the actually observed mergers with the mergers the national government would have mandated if it had chosen a merger pattern directly: centralized mergers would have reduced the number of municipalities to about 650, rather than the 1750 obtained under the decentralized mergers. The improvement over the decentralized pattern would be equivalent to about ¥815 billion of national spending if only the structural component of payoffs is considered.

This final result, however, does not imply that the Japanese government erred in deciding to decentralize the *Heisei* municipal mergers. The merger process is decentralized in order to take advantage of information that is available to neither the national government nor an outside analyst. The resulting mergers will thus have better idiosyncratic characteristics than centrally-planned mergers. If the Japanese government had the option of centralized mergers, but chose a decentralized process instead, then the interpretation of the above results should be that this information problem is economically significant. That is, both centralized and decentralized mergers are substantially worse than the first-best merger pattern that the central government would implement if it had all available information.

The closest related work is Gordon and Knight [2009], who examine school district mergers in Iowa.[4] 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.

[4] Although general models of coalition formation date back at least to von Neumann and Morgenstern [1944], most recent empirical work has focused on a few specific forms of coalition formation games, such as two-sided matching games [Roth, 2008].

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.[5]

The remainder of the paper has the following structure: Section 1 presents the model of local public goods and municipal mergers, Section 2 discusses how the Japanese data can be analyzed using this model, Section 3 describes the estimation strategy, and Section 4 presents the results.

1 Theory

1.1 Public Good Provision

There is a single country, populated by individuals that are distributed across a plane.[6] The location of these individuals is fixed, and they are partitioned into municipalities. For now, suppose that this arrangement of municipalities is also fixed. 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, such as total population. 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

$$q_m c(X_m) = \tau_m Y_m + T_m. \tag{1}$$

The national government obtains funds from an outside source, and spends enough of these on activities that are outside of the model that the marginal opportunity cost

[5]Most empirical studies of political mergers thus far focus on American school districts. Miceli [1993], the earliest example yet found, examines the trade-off that Connecticut school districts faced between efficiencies of scale and locally optimal education quality. Alesina, Baqir, and Hoxby [2004] use a much larger dataset, and examine the relationship between county-level heterogeneity and the number of school districts and other local jurisdictions. While the estimates in each of these papers imply a type of coalition formation game, they do not present an explicit coalition formation model.

[6]This model is based on Greenberg and Weber [1986], Demange [1994], and in particular Alesina and Spolaore [2003].

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

of providing transfers can be treated as constant. Let this marginal cost of funds for the national government be b.

Individual utility is assumed to take the following additively separable form:

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

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 that i is a member of.

The first two terms of this utility function 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 the above equation can be rewritten to treat income as an individual fixed effect:

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

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, and thus choosing θ_m^* is a multidimensional policy decision, a problem which has no single accepted solution concept.

To resolve this, political decision-making at both the local and national level 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 A. At the local level, assume that equal weight is given to all individuals: the local politician acts as a Benthamite social planner. On the other hand, assume that while the national government uses equal weights for individuals within a given municipality, it might have unequal weights across municipalities. Thus, 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, but the national government might choose to provide comparatively large transfers to certain municipalities while giving little to some others.

Specifically, assume that the transfer to municipality m takes the form $T_m = T(X_m, \psi)$ for some parameter vector ψ, and let w_m be the total weight placed on individuals in that municipality by the national government. Then the national government's problem is

$$\max_{\psi} \quad W(T(X,\psi)) \tag{4}$$
$$W(T) = \sum_{m\in M} w_m u_{mm}(T_m) - b\sum_{m\in M} T_m.$$

Here $u_{mm}(T_m)$ corresponds to the max of the local politician's objective function in municipality m, given transfers T_m:[8]

$$u_{mm}(T_m) = \beta_0 \log(1-\tau_m^*) + \beta_1 \log(q_m^* - \beta_3) + \beta_2 \ell_m(\theta_m^*) + \alpha_m + \epsilon_m \tag{5}$$

where

$$\alpha_m = \frac{1}{N_m}\sum_{i\in m}\alpha_i \qquad\qquad \ell_m(\theta) = \frac{1}{N_m}\sum_{i\in m}\ell(i,\theta)$$

and the optimal policies for τ, q and θ are given by

$$\begin{aligned} \tau_m^* &= 1 - \frac{\beta_0}{\beta_0+\beta_1}\frac{Y_m + T_m - \beta_3 c(X_m)}{Y_m} \\ q_m^* &= \frac{\beta_1}{\beta_0+\beta_1}\frac{Y_m + T_m - \beta_3 c(X_m)}{c(X_m)} + \beta_3 \\ \theta_m^* &= \operatorname*{argmin}_{\theta} \ell_m(\theta). \end{aligned} \tag{6}$$

The solution to the national government's problem in (4) will satisfy the first order conditions

$$\frac{\partial W(T(X,\psi))}{\partial \psi_k} = 0 \tag{7}$$

and if there are no restrictions on what transfers the national government can make this simplifies to

$$T_m = \beta_3 c(X_m) - Y_m + \frac{w_m(\beta_0+\beta_1)}{b} \tag{8}$$

for each municipality m.[9] In the special case where all individuals have the same y,

[8] The second m subscript will be used to indicate mergers, considered in the following subsection.

[9] To derive this optimization problem, plug Equations 6 into Equation 3, and drop all the terms that do not include T.

and the government weights all individuals equally, Equation 8 further simplifies to

$$T_m = \beta_3 c(X_m) - aY_m, \tag{9}$$

where $a = 1 - \frac{\beta_0+\beta_1}{yb}$.

Equations 7 - 9 assume that the partition of individuals into municipalities is fixed. If changes such as municipal mergers are possible, then transfer policies satisfying the conditions above may no longer be optimal.

1.2 Municipal Mergers

Consider two possible types of municipal mergers: mergers that are voluntarily agreed upon by the involved municipalities ("decentralized"), and mergers that are mandated by the national government ("centralized"). In some situations the national government may prefer to implement the former, and in others, the latter. This will be illustrated via three simple examples.

The basic setup for both types of mergers is the same.[10] Let $S \subset M$ be a coalition of municipalities that will merge together, and $\mathscr{S} \subset 2^M$ the set of all possible coalitions, including singletons. 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 is $\Pi = \{\pi | \pi \subset \mathscr{S},\ \bigcup \pi = M\}$.

An amalgamated municipality behaves exactly as outlined above in Section 1.1, and is not involved in any further mergers. The utility for individual i in merger S is thus as in Equation 3 above, except replacing m with S:

$$u_i(\tau_S, q_S, \theta_S) = \beta_0 \log(1-\tau_S) + \beta_1 \log(q_S - \beta_3) + \beta_2 \ell(i, \theta_S) + \alpha_i + \epsilon_S. \tag{10}$$

Assume that there is perfect information with the possible exception that the national government may not be able to observe ϵ. Also 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. Finally, assume that the sufficient conditions for a unique (q^*, τ^*, θ^*) political equilibrium, described in Appendix A, hold for all potential mergers. With these assumptions, the post merger choice of q^*_S, τ^*_S, and θ^*_S is

[10] Notation in this section is based on that used by Banerjee, Konishi, and Sönmez [2001].

known in advance for every potential coalition S, and is exactly as in Equation 6, except replacing municipality m with coalition S.

The payoff for municipality m participating in coalition S can be described as in Equation 5:[11]

$$u_{mS} = \beta_0 \log(1 - \tau_S^*) + \beta_1 \log(q_S^* - \beta_3) + \beta_2 \ell_m(\theta_S^*) + \alpha_m + \epsilon_S. \tag{11}$$

Now consider the difference in payoffs for municipality m of participating in coalition S versus remaining a singleton:[12]

$$\begin{aligned} u_{mS} - u_{mm} = &\beta_0(\log(1 - \tau_S^*) - \log(1 - \tau_m^*)) \\ &+ \beta_1(\log(q_S^* - \beta_3) - \log(q_m^* - \beta_3)) \\ &+ \beta_2(\ell_m(\theta_S^*) - \ell_m(\theta_m^*)) \\ &+ \epsilon_S - \epsilon_m, \end{aligned} \tag{12}$$

The payoffs to municipality m considered in Equation 12 depend only on the characteristics of m and S, and a transfer policy that the national government has already committed to.[13] Municipal mergers are thus being treated as a pure hedonic coalition formation game, where the payoff to each player depends only on the coalition to which they belong, and not on what other coalitions occur.[14]

First, consider the case where mergers are effected via a decentralized process,

[11]The first term in Equation 5 is the utility received by individuals in municipality m from their private consumption. Due to the log functional form, the income term itself becomes the fixed effect α_m, following Equation 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. The third term, however, takes into account distance to the location where the public service is provided, and this may differ substantially between municipalities in S. For example, if residents of m would be close to θ_S^*, while residents of m' would be far away, then the disutility from distance will be less severe for m than for m' if coalition S occurs.

[12]For simplicity, the payoff to the singleton merger $\{m\}$ is denoted by u_{mm}, following Equation 5, rather than $u_{m\{m\}}$.

[13]The first term is positive if the tax rate chosen in the merger is lower: $\frac{Y_S+T_S-\beta_3 c(X_S)}{Y_S} > \frac{Y_m+T_m-\beta_3 c(X_m)}{Y_m}$. The second term is positive if the quality of the public good provided is higher: $\frac{Y_S+T_S-\beta_3 c(X_S)}{c(X_S)} > \frac{Y_m+T_m-\beta_3 c(X_m)}{c(X_m)}$. The third term (the difference in disutility from distance) will be zero or negative so long as distance is undesirable ($\beta_2 < 0$), because θ_m^* minimizes ℓ_m.

[14]This is the game introduced by Dreze and Greenberg [1980], except without the possibility of even within-coalition transfers.

where a merger requires approval from all participating municipalities, and assume that this decentralized decision-making involves each municipality making decisions based on the utility function in Equation 5.[15] Here, the inability for municipalities to negotiate transfers may prevent some coalitions from forming. To see this, rewrite Equation 5 as

$$u_{mS} = v_{mS} + \epsilon_S \tag{13}$$

and consider the following example:

Example 1. *Suppose that there are two municipalities with identical characteristics, the cost function c has constant efficiencies of scale, and the only possible merger is* $S = \{m, m'\}$ *with* $v_{mm} = v_{mS} = v$*. Then if* $\epsilon_m > 0$*,* $\epsilon_{m'} < 0$*,* $|\epsilon_m| > |\epsilon_{m'}|$*, the merger will not occur.*

If transfers between municipalities were possible, and u did not exhibit too much curvature in private consumption, then m would offer a transfer to m' in exchange for m' agreeing to the merger, and the merger would occur. However, the payoffs in Equation 12 do not allow for this sort of inter-municipality payment to be negotiated.

The national government thus might wish to mandate mergers instead of allowing decentralized mergers. If the national government has perfect information then it will mandate that this merger occur when

$$(u_{mS} - u_{mm})w_m + (u_{m'S} - u_{m'm'})w_{m'} > 0 \tag{14}$$

if the national government's transfer policy is as given in Equation 8.

On the other hand, if the national government does not know ϵ then it may choose to implement decentralized mergers instead of mandating a certain pattern of municipal mergers:

Example 2. *If, in the case described in Example 1, the national government weights all individuals equally, is not restricted in the transfers it can make, and does not observe* ϵ*, then if* $E[\epsilon_m - \epsilon_S] = E[\epsilon_{m'} - \epsilon_S] = 0$ *the national government will choose to implement a decentralized merger policy.*

[15] As discussed below, with this sort of decentralized system, once there are multiple potential mergers involving some of the same municipalities it may no longer be obvious which mergers will occur. For now, however, consider the case where each municipality can participate in only one merger.

A centralized merger policy would have expected payoff of v, regardless of whether the merger is mandated or prohibited, because with the optimal transfer scheme in this case there is no difference in total transfers. The decentralized policy, on the other hand, will result in a merger when both $\epsilon_S - \epsilon_m > 0$ and $\epsilon_S - \epsilon_{m'} > 0$, and no merger otherwise. This improves on either centralized policy choice, and thus the national government will choose to implement decentralized mergers even if it had the option of controlling mergers centrally.

If the government has decided to implement a decentralized merger policy, then the optimal transfer policy may not be the same as that given in the preceding section:

Example 3. *Suppose that the situation is as described in Example 2, except now suppose that c exhibits efficiencies of scale. Then if ϵ has full support, the transfer policy in Equation 8 is not optimal for decentralized mergers.*

To see this, let $p(T)$ be the probability that the merger will occur if the government chooses transfer vector T. As in Equation 13, let $v_{mS}(T)$ be the non-idiosyncratic component of $u_{mS}(T)$. The national government's problem is

$$\begin{aligned}\max_{T} \quad & p(T)(v_{mS}(T) + \mathrm{E}[\epsilon_S|\pi^* = \{S\}, T] - bT_S) \\ & +(1-p(T))(v_{mm}(T) + \mathrm{E}[\epsilon_m|\pi^* = \{\{m\},\{m'\}\}, T] - 2bT_m).\end{aligned} \tag{15}$$

where π^* is the partition that is actually observed, which depends on T as well as ϵ. Increasing T_S helps to resolve two externalities, as local governments consider neither their partners' payoffs nor the national budget when making merger decisions. Thus, the transfer policy that was optimal for the national government when municipal boundaries were fixed is no longer optimal when there is the possibility of municipal mergers.[16]

[16]Start with the transfer policy given by Equation 8, and consider a small deviation that increases transfers to the municipalities if they merge and decreases transfers by an equivalent total amount if they do not. The cost of this deviation is that the transfer policy is no longer optimal given fixed boundaries. This cost is second-order. On the other hand, there are two first-order benefits. First, because c exhibits efficiencies of scale then the national government spends less money in expectation because $T(X_S) < 2T(X_m)$ for the policy given by Equation 8. Second, in the case where one municipality prefers the merger but the other is indifferent, the indifferent municipality does not internalize the benefits of the merger to its partner, but could be encouraged to merge via a higher T_S.

1.3 Solution Concept

While in the examples presented thus far it has been intuitively clear for any given ϵ which mergers would occur if the process were decentralized, unfortunately there are other situations where the decentralized outcome is not so obvious. Consider, for example, the classic "roommates problem":

Example 4 (Gale and Shapley 1962)**.** *Suppose* $M = \{1, 2, 3\}$ *and preferences are*

$$
\begin{aligned}
&\{1,2,3\} \prec_1 \{1\} \prec_1 \{1,3\} \prec_1 \{1,2\}, \\
&\{1,2,3\} \prec_2 \{2\} \prec_2 \{1,2\} \prec_2 \{2,3\}, \\
&\{1,2,3\} \prec_3 \{3\} \prec_3 \{2,3\} \prec_3 \{1,3\}.
\end{aligned}
\tag{16}
$$

Given these preferences, there is no intuitively 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.

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 4.

The environment considered in this paper is simpler than that considered in Ray and Vohra [1997], and thus a simpler solution concept 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.*

Proof. See Appendix B. □

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 4 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. Theorem 1 thus rules out partitions that should definitely not occur, but does not specify precisely what partition will occur.[17]

2 Japanese Context

2.1 Public Good Provision

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 administration.[18] In order to apply the theoretical model presented in Section 1 to these municipalities, a number of simplifying assumptions are necessary. In particular, terms in the municipal budget constraint (Equation 1) and the individual utility function (Equation 3) need to be matched to actually observed data. Begin by considering the municipal budget, as it constrains the first two terms of the individual utility function.

The three elements of Equation 1 are expenditures on local services, transfers from the national government, and local taxes. Beginning in the 1950s, the Japanese government established a "national standard" 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".[19] In the period before the *Heisei* municipal mergers began, transfers to municipalities were determined by a formula

[17] The approach used by Gordon and Knight [2009] to guarantee uniqueness is not possible in this case because the $\ell(i, \theta_m)$ in Equation 3 generates the possibility of preference cycles as in Example 4. The non-cooperative approach used by Diermeier, Eraslan, and Merlo [2003] requires additional assumptions regarding the exact process by which coalitions were formed.

[18] Public goods that 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.

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

quite similar to Equation 9:

$$T_m = \max(\tilde{c}(X_m) - .75\bar{\tau}Y_m, 0) \tag{17}$$

where $\tilde{c}(X_m)$ is the national government's estimate of the cost to a municipality with characteristics X_m of providing public goods at the "national standard" quality, and $\bar{\tau}Y_m$ the amount of tax revenue the municipality should be able to collect if taxes are charged at the "standard" rate. The cost estimate $\tilde{c}$ is sometimes referred to as the "Standard Fiscal Need", and is calculated by a formula equivalent to

$$\tilde{c}(X_m) = \sum_{k=1}^{24} X_{mk} \cdot \bar{c}_k(1 + \tilde{H}_k(X_m)) + \zeta_m. \tag{18}$$

Here the public good is viewed as a sum of 24 component goods, such as firefighting, care for the elderly, resident registration, and so forth. Each of these component goods is associated with a quantity measure X_{mk} and an estimated unit cost $\bar{c}_k$, and an adjustment coefficient $\tilde{H}_k$. The adjustment coefficient is created by multiplying and adding together a set of (usually) decreasing splines determined by X_m. These adjustments generally resulted in higher per capita cost estimates for smaller municipalities, as in the case of the *dankai* ("step") adjustment shown in Figure 4. There are also some additional costs ζ_m that are not generally subject to adjustment. Further details regarding Equation 18 are provided in Appendix C.

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 $q_m^f = 5$ firefighting, 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

$$c(X_m) = \sum_{k=1}^{24} X_{mk} \cdot \bar{c}_k \cdot (1 + H_k(X_m)) + \zeta_m \tag{19}$$

and assume that any bias in the Japanese government's estimate of $\tilde{c}$ can be described by parameters ψ in the following way:

$$\tilde{c}(X_m) = \sum_{k=1}^{24} X_{mk} \cdot \bar{c}_k \cdot (\psi_0 + \psi_1 H_k(X_m)) + \zeta_m. \tag{20}$$

As there are no natural units for quality, use the normalization $\psi_0 = 1$ for the Japanese's government's estimates prior to the merger period.[20]

Fitting actual municipal taxes to the model in Section 1 requires the same sort of assumptions as for service provision. The principal source of local tax revenues is property tax, but there are multiple tax bases each of which can be taxed at a different tax rate, while the model in Section 1 has only a single tax base Y with a single tax rate τ. To resolve this issue, use the same approach as for the cost of public services: assume that municipalities must tax each tax base in direct proportion to the national government "standard" tax rates on these tax bases. This appears to be a reasonable assumption, as tax rates do not differ much from the standard rates.

An additional issue with taxation is that there is some bureaucratic imprecision regarding the $\bar{\tau} Y_m$ term in Equation 17, sometimes referred to as "Standard Fiscal Revenue". 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. Issues regarding taxation are discussed in more detail in Appendix C. Below, 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 municipal budget constraint consists of the three terms just discussed: taxes, transfers, and expenditures. The municipal tax base and national government transfers are available from published sources, described further in Appendix C. The cost of providing the public good, however, is unknown and will be estimated as described in the following section. For a given cost function, and for given parameters β, the τ_m^* and q_m^* that will be selected by each municipality can be calculated using the formulae from Equation 6. The first two terms of the individual utility function (Equation 3) can then be calculated for any individual and municipality.[21]

The remaining structural term in Equation 3 is distance, ℓ. This is taken to be

[20] Here ψ_1 is assumed to be the same for all component goods, which corresponds to the national government's estimates of efficiencies of scale $\tilde{H}$ in Equation 18 having the form $\tilde{H}_k(X_m) = \psi_1 H_k(X_m)$.

[21] The individual fixed effect α_i can also be calculated, but because the data used involves choices between coalitions, this term will always disappear, as shown in Equation 12.

geographic distance, and is calculated using census grid square data on population.[22] 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.[23] Surveys conducted around the time of the recent mergers reveal a widespread perception that there were unexploited efficiencies of scale in the provision of these services.[24] It was generally understood that a municipal merger involving a smaller municipality and a larger one would result in the closure of city hall and some other facilities in the smaller municipality, and that this would result in substantial cost savings. This lead to a concern among residents of smaller municipalities that after a merger public services would only be provided at a location much further away than was the case currently. As an increase in geographic distance to public facilities was perceived as a major cost of a municipal merger, it seems appropriate to use geography as the metric for heterogeneity in the structural part of the model.[25]

2.2 Municipal Mergers

If a coalition S of municipalities decided to amalgamate, transfers would in principle be calculated as in Equation 17:[26]

$$T_S = \max(\tilde{c}(X_S) - .75\bar{\tau}Y_S, 0) \tag{21}$$

which would result in savings for the national government, as $\tilde{H}$ in Equation 18 usually decreased with size. Thus, local residents would oppose mergers that the national government would like to see occur. The exact behaviour of residents was

[22]These calculations are described in more detail in Appendix C.

[23]Although many services at city hall could be accessed via mail, telephone, or the internet, it is common and in some cases required to visit in person.

[24]The most popular response to questions regarding the potential benefits of municipal mergers was “avoid duplication of facilities / avoid useless capital expenditures” in Kyoto, “reduce expenditures by improving administrative efficiency, eliminating duplicate facilities, and reducing personnel” in Yamanashi prefecture, “reduce personnel expenses” in Akita, and “reduce personnel and other expenditures and improve efficiency” in Okinawa. Unfortunately, these surveys are difficult to analyze quantitatively, as they were only conducted in a few prefectures, and a different questionnaire was used in each prefecture.

[25]A 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.

[26]Some merger incentives were offered, as described below.

determined in part by the relationship between the true cost function c and the national government's estimate $\tilde{c}$.

During the fiscal difficulties of the early 1990s, the Japanese national government implemented a series of reforms designed to reduce the total transfers provided to municipalities while attempting to minimize the negative effects of this decrease. First, the government substantially reduced transfers, particularly to the smallest municipalities. This was effected mainly by replacing $\tilde{H}_k$ with $\tilde{H}_k^{\text{new}}$, which was less generous towards smaller municipalities as is shown in Figure 5.[27] These new transfers T^{new} provided smaller municipalities with an incentive to merge so as to avoid having to either sharply reduce the quality of service that they provided to their residents, or increase the tax rate charged.

Second, with these new transfers in place, the government then allowed municipalities to keep more of the savings from a merger. The "merger" formula

$$T_S^{\text{merger}} = \sum_{m \in S} T_m^{\text{new}} \tag{22}$$

had previously been used for up to five years following a merger, but in 1999 this was increased to ten years starting from the date of the merger.[28] This incentive began to be phased out in 2006, which motivated many municipalities to finalize mergers in 2005.[29]

A final incentive for mergers was the *Gappei Tokureisai*, special subsidized bond issues allowed for municipalities planning amalgamation.[30] The value of these bonds

[27] Although this change took place around 2003, it was announced earlier.

[28] An intermediate amount between T_S^{new} and T_S^{merger} was offered for years 11-15 following a merger.

[29] By 2006 there were only 1,844 municipalities remaining, down from 3,255 at the start of the merger period. The situation in Shizuoka Prefecture at that time is shown in Figure 6. A small number of mergers occurred during the phase-out period, reducing the final number of municipalities to 1,750 in 2010; for the purposes of this paper, these mergers are treated as though they were finalized prior to 2006, and implementation was simply delayed for exogenous reasons. Explaining 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.

[30] The official explanation for these 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. Relative to the incentive provided by the switch from $\tilde{c}_m$ to $\tilde{c}_m^{\text{new}}$ in the calculation of transfers, these bonds have a relatively small effect on incentives to merge, and thus for simplicity the the bonds are treated entirely as an additional

is calculated based on the subsidy offered, using information from Ishihara [2000]. Municipalities are presumed to able to save in order to equalize the quality of public services and the municipal tax rate between the decade immediately following the merger, when incentives are provided, and following decades. As the rules for all the financial incentives for mergers are known, for each coalition S the optimal tax rate and public good quality can be calculated for any given cost function and parameters β.[31]

Although a large number of mergers occurred overall, very few of these mergers involved municipalities in the most metropolitan prefectures. Define a prefecture as "metropolitan" if fewer than 10% of its municipalities have a population of less than 10000, and define a prefecture as "rural" if more than 65% of its municipalities have a population of less than 10000. Table 1 shows summary statistics for municipalities in different classes of prefectures. Municipalities in metropolitan prefectures were much less likely to merge during the merger period than those in the other sorts of prefectures.

3 Estimation

There are four parameters of interest from Equation 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). An additional parameter of interest is the relationship between the Japanese government estimates of efficiencies of scale and the true efficiencies of scale: this is ψ_1 in Equation 20. One hypothesis of particular interest is that there are in reality no efficiencies of scale, but this corresponds to $\psi_1 \to \infty$ and is thus difficult to test using the parameterization of Equation 20. Thus, define $\beta_4 = \frac{1}{\psi_1}$.[32] The case where there are no efficiencies of scale in the production of public goods then corresponds to $\beta_4 = 0$, and the case where the national government's estimates are correct corresponds to $\beta_4 = 1$.

These β parameters can be estimated by examining the mergers that actually occurred in Japan and comparing them to ones that could have occurred but did not,

incentive, with the direct financial costs of merging ignored.

[31] An overview of the rules described in the "Special Municipal Merger Law" is described in Appendix A.1.1.

[32] This then implies that $H_k(X_m) = \beta_4 \tilde{H}_k(X_m)$, which is more convenient from an empirical perspective, because $\tilde{H}_k$ is known while H_k is not.

using the data on national government transfers and efficiencies of scale described in Section 2 and further explained in Appendix C. To do this, first rewrite the utility function in Equation 13 to make explicit the fact that the values of the β parameters affect the structural component, but do not affect the idiosyncratic component:

$$u_{mS}(\beta) = v_{mS}(\beta) + \epsilon_S. \tag{23}$$

Assume that ϵ is distributed normally, with the distribution of ϵ_S identical to that of $\epsilon_{S'}$, but not necessarily independent. Furthermore, note that, as is standard in discrete choice models, multiplying u by a positive constant has no effect on preferences. Thus, as a normalization, multiply such that $\epsilon_S \sim N(0,1)$. Let β^0 be the true value of β. Estimation will be based on four types of moment inequalities:

1. At β^0, it should be possible to find values of the idiosyncratic shocks that rationalize the observed mergers, and are not “too extreme” relative to the $N(0,1)$ distribution from which they are assumed to have been drawn.

2. At β^0, if the national government had not implemented the merger promotion policies, the number of mergers that would have occurred should not be “too high” relative to the number of mergers that occurred prior to 1999.

3. At β^0, the number of mergers actually observed in metropolitan prefectures should not be “too low” relative to the number of mergers predicted in these prefectures by the model.

4. At β^0, the tax rates actually charged by municipalities should be “similar” to those that the model predicts should be charged.

The technical definitions for terms in quotation marks will be given below. The first two types of moment inequalities should hold for metropolitan prefectures, mixed prefectures, and rural prefectures, as defined at the end of Section 2. The remainder of this section has the following form: first, the covariance of ϵ_S and $\epsilon_{S'}$ is discussed, and then the details of each of the above four types of moment inequalities are presented. Finally, confidence sets for the estimated parameters are calculated.

3.1 Structure of Idiosyncratic Shocks

There are three important points regarding the set $\mathscr{S}$ of all potential mergers. First, it is potentially very large, containing up to 2^M elements. Second, it is not always clear which coalitions should be in this set: 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. Finally, it is implausible that ϵ is i.i.d across different coalitions: if $S = \{m_1, m_2, ..., m_{14}, m_{15}\}$, and $S' = \{m_1, m_2, ..., m_{14}\}$, then a reasonable econometric model should have ϵ_S correlated with $\epsilon_{S'}$. The following construction makes it possible to generate shocks that are $\epsilon_S \sim N(0,1)$, not independent but identically distributed, and to consider the shocks associated with some coalitions without enumerating all potential coalitions. The basic assumption comes from the literature on ethnic fragmentation: under certain conditions, heterogenous jurisdictions produce bad results for all residents, not only those far from the median voter. While Japan is not known for extreme ethnic or linguistic heterogeneity, one could imagine even minor cultural differences playing such a role.[33]

First, suppose that each individual resident makes 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 be

$$\bar{\omega}_m = \frac{1}{N_m} \sum_{i=1}^{N_m} \omega_i \tag{24}$$

$$s_m^2 = \frac{1}{N_m - 1} \sum_{i=1}^{N_m} (\omega_i - \bar{\omega}_m)^2, \tag{25}$$

because there are N_m residents making i.i.d. draws. Within-municipality heterogeneity is captured by the sample variance, so let

$$\epsilon_m = -f(X_m) \log s_m^2, \tag{26}$$

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.

Defining $\bar{\omega}_S$ and s_S^2 in the same way for any coalition S, the standard relationship

[33] Costa and Kahn [2003] provides a discussion of some of these issues.

for sample variances will hold:

$$s_S^2 = \frac{\sum_{m\in S}(N_m - 1)s_m^2 + \sum_{m\in S}\sum_{m'\in S}\frac{N_m N_{m'}(\bar{\omega}_m - \bar{\omega}_{m'})^2}{N_S}}{N_S - 1} \tag{27}$$

Now define the vectors $\bar{\omega}_M$ and s_M to be the sample means and standard deviations for all municipalities. It is possible to calculate ϵ_S for any coalition S given only $\bar{\omega}_M$ and s_M. Let $\epsilon(\bar{\omega}_M, s_M)$ be the vector resulting from this calculation. The elements of both $\bar{\omega}_M$ and s_M have known distributions, which will be helpful when computing moment inequalities.

With this construction of ϵ, for any guess $\hat{\beta}$ for the parameter vector, any observed partition can be rationalized: simply choose s_S^2 sufficiently close to zero if S is in the observed partition, and large otherwise, and then choose $\bar{\omega}_S$ and $\bar{\omega}_{S'}$ such that $(\bar{\omega}_S - \bar{\omega}_{S'})^2$ is sufficiently large to discourage any coarsenings into larger coalitions.

Finally, using the approximation

$$\begin{aligned}\log s_S^2 &\simeq \log(1+\delta_S), \qquad \delta_S \sim N(0, \frac{2}{N_S - 1})\\ &\simeq \delta_S\end{aligned}$$

it is the case that if $f(X_S) = \sqrt{\frac{N_S - 1}{2}}$, then $\epsilon_S \sim N(0,1)$, n.i.i.d, as desired.

3.2 Moment Inequalities

The first moment inequality used is based on rationalizing the observed coalition structure. First choose an arbitrary function $h(\bar{\omega}, s|X)$.[34] Then define h^* as

$$h^*(\pi|X,\beta) = \min_{\bar{\omega}_M, s_M} h(\bar{\omega}_M, s_M|X) \quad \text{s.t. } \pi \in \Pi^*(\epsilon(\bar{\omega}_M, s_M)|X,\beta) \tag{28}$$

where Π^* is the stable set evaluated at parameters β, and with idiosyncratic shocks ϵ as determined by $\bar{\omega}_M$ and s_M. Thus, h^* is a lower bound for h given that partition π was observed, and that parameter values are β. It is always the case that

$$h(\bar{\omega}_M^0, s_M^0|X) \geq h^*(\pi^0|X,\beta^0) \tag{29}$$

[34] Good choices for h seem to be functions that will give high values when $\bar{\omega}$ and s are extreme relative to the distributions from which they were assumed to have been drawn.

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, and thus $(\bar{\omega}_M^0, s_M^0)$ could be chosen in Equation 28. Consider the moment

$$g_1(\pi, \beta|X) = Eh(\bar{\omega}_M, s_M|X) - h^*(\pi|X, \beta). \tag{30}$$

This will be positive in expectation at the true parameter value β^0, because

$$\begin{aligned} Eg_1(\pi, \beta^0|X) &= Eh(\bar{\omega}_M, s_M|X) - Eh^*(\pi, \beta^0) \\ &\geq 0 \end{aligned} \tag{31}$$

because Inequality 29 holds at every realization of $(\bar{\omega}_M, s_M)$ and π, and thus also holds in expectation. Computation is explained in more detail in Appendix D.1, and a very simple example is given in Appendix D.4.

The second 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. The moment

$$g_2(Q, \beta|X) = Q - \mu_Q^*(\beta|X) \tag{32}$$

can then be used as a moment inequality, where Q is the number of municipalities involved in mergers in the 1979-1998 period. Computation is explained in more detail in Appendix D.2.

The third moment inequality used is based on the column in Table 1 showing that very few mergers occurred in “metropolitan” prefectures, defined in this paper as prefectures with fewer than 10% of municipalities having a population of less than

10000. The same argument used in the previous subsection can thus be extended to mergers actually observed during the merger period: given the national government's actually implemented policies, the number of mergers observed should not be anomalously low.

Specifically, let Q^{99} be the number of municipalities actually participating in mergers in the 1999-2010 period. Then the moment $g_3(Q, \beta|X) = Q - \mu^*_Q(\beta|X)$, where F^*_Q is as defined in the previous section, can be used as a moment inequality because

$$\begin{aligned} E_Q g_3(Q^{99}, \beta^0|X) &= \mu^{99}_Q - \mu^{*99}_Q(\beta^0|X) \\ &\geq 0 \end{aligned} \tag{33}$$

due to stochastic dominance and Q^{99} having been drawn from F_Q.

The fourth and final moment inequality used relies on the fact that tax rates that are actually charged 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. One 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. Specifically, suppose that the observed tax rates are a function of optimal tax rates plus some noise:

$$\tau^{**}_m = \max(\tau^*_m(\beta) + \varepsilon_m, \bar{\tau}), \tag{34}$$

where τ^*_m is taken from Equation 6. 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

$$\tau^{**}_m = \max(\tau^*_m(\beta) + \gamma X_{mk} + \varepsilon_m, \bar{\tau}), \tag{35}$$

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

4 Results

Results are shown in Table 2. Since mergers do not cross prefectural boundaries, each prefecture is treated as a separate coalition formation game, and asymptotics are with respect to the number of prefectures.[35] Although the model is only set identified in theory, the results show that the minimizer of the test statistic is a single point. This result is standard in the literature.[36]

This value of β_2 roughly implies that an individual would be willing to have a municipal policy that was 1km more distant in exchange for about ¥3500 per year.[37] Using this value of β_2, if the population of Japan were uniformly distributed across the country, and a social planner could set entirely new boundaries for municipalities, then the optimum size for a municipality would be

$$\beta_0 \log(\frac{yN_m - \beta_3 c(N_m)}{yN_m}) + \beta_1 \log(\frac{y_m N_m - \beta_3 c(N_m)}{c(N_m)}) + 0.377\beta_2\sqrt{N_m/340}, \qquad (36)$$

where 0.377 is a coefficient for the average distance to the centroid based on hexagonal packing, and $P_m/340$ the area in square kilometres given the population density of Japan (340 per km^2). This formula yields an optimal municipal population of about 200,000. This suggests 635 municipalities for all of Japan, compared to the 1750 actually present at the end of the merger period.

The estimate for β_3 indicates that the view of the government estimates of the cost of providing public goods as an estimate of the cost of providing the "national minimum" 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 (that is, the government cost estimates are the estimate for the minimum possible public expenditure), but is statistically different from zero.

On the other hand, β_4, the degree to which the central government's estimates of efficiencies of scale in the provision of public services match the true efficiencies of scale, is estimated to be about 0.5. This means that two null hypothesis can be

[35]There 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.

[36]When 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. The models used to produce the estimates in Table 2 are not rejected at the 95% level.

[37]This is not statistically different from the estimate of ¥1000 obtained in Hayashi, Nishikawa, and Weese [2010], based on stated preference data regarding potential municipal mergers.

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 initial government estimates ($\beta_4 = 1$).

4.1 Inefficiency and Alternative Policies

To calculate a measure of inefficiency, it is necessary to determine the national government's objective function. Assume that weights per capita are equal for all individuals in the same prefecture, and differences between prefectures are given by the formula

$$w_m N_m = \phi_0 + \phi_1 \text{seats per capita}_m \tag{37}$$

where "seats per capita" is the legislative representation allocated to the prefecture of municipality m in the upper house of the national Diet.[38]

Next, assume that before the merger period the national government was maximizing its objective (4) under the assumption that boundaries were fixed, and given its weights on individuals. That is, the national government's estimate of the cost function $\tilde{c}$, which determines transfers through Equation 17 and is parameterized by ψ as shown in Equation 20, was the argmax of W in Equation 4.

Using the $\hat{\beta}$ from Column I of Table 2, $\hat{\phi}_0$ and $\hat{\phi}_1$ can be estimated via GMM.[39] The estimate for the intercept, $\hat{\phi}_0 = -0.04$, is not statistically significant, while the slope $\hat{\phi}_1 = 0.18$ is. Thus, the hypothesis that the national government weights all individual equally is rejected, while the hypothesis that the national government weights individuals based on their legislative representation is not rejected. Now, using these $\hat{\phi}$ estimates, some simple counterfactual analysis can be performed.

One parameter that is unknown is the cost of public funds during the merger period. The assumption in this paper is that the national government's policy change

[38]This is an extremely simple measure of legislative malapportionment, as it ignores the lower house entirely. Lower house malapportionment is more complex, and varies within prefecture. The model presented in Section 1 does not work well with weights that differ within the same prefecture, because this implies that for some municipal mergers, the national government would want different local policies than the local government will implement, as the national government has different weights for different individuals within the same (amalgamated) municipality. Thus, only upper house malapportionment is used, as representation is at the prefectural level in the upper house.

[39]The "tax floor" specification is preferred both because it appears to fit anecdotal evidence better, and because in counterfactual policy exercises it prevents the national government from using the transfer system to massively redistribute income from richer to poorer municipalities via negative municipal tax rates.

was prompted by an increase in this cost from $b = 1$ to some higher cost $b' > b$. Intuitively, the lowest benefits for mergers will be when $b' \simeq b$, and higher values of b will lead to larger benefits.[40] Thus, assume the scenario that $b' = b$.[41]

To determine the benefit to the national government of the observed mergers occurring, take the actual transfer policy used in the merger period and calculate a lower bound for W from Equation 4 using weights $\hat{w}$ calculated using $\hat{\phi}$ and Equation 37. An exact calculation is not possible, because W depends on u, which includes ϵ, which is unobserved. The point of decentralizing the mergers, however, is to obtain a configuration that has good unobserved characteristics. Thus

$$\begin{aligned} W(T^{\text{merger}}) &> V(T^{\text{merger}}) - b' \sum_{S \in \pi^0} T_S^{\text{merger}} \\ V(T^{\text{merger}}) &= \sum_{S \in \pi^0} \hat{w}_S v_{mS}(T_S^{\text{merger}}). \end{aligned} \tag{38}$$

In comparison to the national government's objective function under the original municipal boundaries, the mergers provide a benefit equivalent to about ¥160 billion of public funds. This is likely a substantial underestimate of the benefits, as it ignores idiosyncratic benefits. An alternative estimate, assuming that idiosyncratic terms were sufficiently positive such that all municipalities participating in a merger preferred that to remaining a singleton, gives an estimate about twice as large: ¥335 billion.

Now, suppose that, rather than relying on decentralized mergers, the government had simply imposed the very boundaries that were actually observed post-merger. The government could then have implemented a transfer policy very different than the one it actually chose, because there are no longer any incentive problems at the municipal level regarding potential mergers. In the case where mergers were mandated, the optimal transfer policy would be $\psi_0' = 1.03$, $\psi_1' = 1.09$, which is very close to the original transfer policy. Comparing this policy to the actual transfer specified

[40] An upper bound for b' can be calculated by assuming, following Example 3, that a cut in transfers to non-merging municipalities would not result in a worse merger pattern. Using $\hat{\phi}$, and assuming that this cut would have to take the form given in Equation 20, this bound can be calculated as $b' < 1.4$. For $b' \in [1, 1.4]$, the lowest benefits for the observed mergers are indeed at $b' = 1$.

[41] One question here is why, if the cost of public funds did not change, the calculated benefits for mergers are positive. Municipal borders in Japan had remained mostly unchanged since the 1960s, and even if the borders set in the 1960s were optimal, benefits to realignment were likely appeared over time. If there are advantages to large one-shot rearrangements, rather than piecewise modifications, then there is an option value to not implementing a municipal merger policy in any given year. Positive benefits are thus consistent with the policy not being implemented.

during the mergers, there a benefit to the national government equivalent to ¥250 billion in public funds.

Finally, consider the case where the government chooses a merger pattern centrally. For simplicity, suppose that the government chose the transfer policy that it actually implemented, even though this would be suboptimal. Even with this additional assumption, finding the optimal pattern of mergers is a difficult combinatorial problem. However, a relatively simple "greedy" algorithm yields a lower bound.[42] This bound is a benefit equivalent to ¥1150 billion in additional spending. If the actually implemented mergers provided a benefit equivalent to ¥335 billion, then the optimal merger pattern would have provided the equivalent of an additional ¥815 billion. This suggests that either the information problems faced by the national government were very severe, or centralized mergers were not a feasible option for other reasons. The optimal pattern calculated for these centralized mergers has a final total of 645 municipalities; however, unlike the calculated benefit this is not a bound, and thus the true optimal pattern could involve more or fewer municipalities. This number is very similar to the 635 municipalities calculated above under the assumption of uniform population distribution and hexagonal packing, however, suggesting that other approximation algorithms would likely yield similar results.

5 Conclusion

This paper examined the inefficiency arising from local decisions over boundary changes, using Japanese data and an estimator based on moment inequalities. Estimation is feasible due to a special covariance structure assumed for the error term.

Municipalities take into account neither the benefits of a merger to their merger partners, nor its effect on national spending. The national government thus chooses a second-best transfer policy that provides less equalization but greater incentives for municipalities to participate in mergers. The resulting merger pattern is also second-best, with smaller mergers than the national government would have preferred. If the national government had centralized the merger process, estimates suggest that the post-merger political configuration would have had only about 650 municipalities, rather than the approximately 1750 actually observed.

[42] Start by merging the pair of municipalities where the merger would produce the greatest benefit. Continue until there are no more mergers that will improve the objective function.

References

Daron Acemoglu. Why not a political coase theorem? social conflict, commitment, and politics. *Journal of Comparative Economics*, 31(4):620–652, December 2003.

Alberto Alesina and Enrico Spolaore. On the number and size of nations. *The Quarterly Journal of Economics*, 112(4):1027–1056, November 1997.

Alberto Alesina and Enrico Spolaore. *The size of nations*. MIT Press, Cambridge, Mass, 2003. ISBN 0262012049.

Alberto Alesina, Reza Baqir, and Caroline Hoxby. Political jurisdictions in heterogeneous communities. *The Journal of Political Economy*, 112(2):348–396, April 2004.

Donald W. K. Andrews and Patrik Guggenberger. Validity of subsampling and "Plug-In asymptotic" inference for parameters defined by moment inequalities. *Econometric Theory*, 25(03):669–709, 2009.

Suryapratim Banerjee, Hideo Konishi, and Tayfun Sönmez. Core in a simple coalition formation game. *Social Choice and Welfare*, 18(1):135–153, January 2001.

Jeffrey Banks and John Duggan. Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates. In David Austen-Smith and John Duggan, editors, *Social Choice and Strategic Decisions*, Studies in Choice and Welfare, pages 15–56. Springer Berlin Heidelberg, 2005.

S. Barberà and A. Gerber. A note on the impossibility of a satisfactory concept of stability for coalition formation games. *Economics Letters*, 95(1):85–90, April 2007.

Anna Bogomolnaia and Matthew O. Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, February 2002.

Federico Ciliberto and Elie Tamer. Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828, November 2009.

Dora L. Costa and Matthew E. Kahn. Civic engagement and community heterogeneity: An economist's perspective. *Perspectives on Politics*, 1(01):103–111, 2003.

Gabrielle Demange. Intermediate preferences and stable coalition structures. *Journal of Mathematical Economics*, 23(1):45–58, January 1994. ISSN 0304-4068.

Andrew DeWit. Dry rot: The corruption of general subsidies in japan. *Journal of the Asia Pacific Economy*, 7(3):355–378, October 2002.

Effrosyni Diamantoudi and Licun Xue. Coalitions, agreements and efficiency. *Journal of Economic Theory*, 136(1):105–125, September 2007.

Daniel Diermeier, Hulya Eraslan, and Antonio Merlo. A structural model of government formation. *Econometrica*, 71(1):27–70, 2003.

J. H. Dreze and J. Greenberg. Hedonic coalitions: Optimality and stability. *Econometrica*, 48(4):987–1003, May 1980.

Jeremy T. Fox. Estimating matching games with transfers. Technical report, NBER, October 2008.

D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, January 1962.

Nora Gordon and Brian Knight. A spatial merger estimator with an application to school district consolidation. *Journal of Public Economics*, 93(5-6):752–765, June 2009.

Joseph Greenberg. Stability when mobility is restricted by the existing coalition structure. *Journal of Economic Theory*, 21(2):213–221, October 1979.

Joseph Greenberg and Shlomo Weber. Strong tiebout equilibrium under restricted preferences domain. *Journal of Economic Theory*, 38(1):101–117, February 1986.

Masayoshi Hayashi, Masashi Nishikawa, and Eric Weese. Inefficient municipal boundaries. Technical report, Aoyama Gakuin, 2010. [In Japanese].

Nobuo Ishihara. *Shin Chihou Zaisei Chousei Seido Ron*. Gyousei, Tokyo, new edition, 2000. [”New Local Finance Regulation System”].

Akihiko Kawaura. Self-Serving mayors and local government consolidations in japan. Technical Report 10-14, 2010.

Ehlers Lars. Von Neumann–Morgenstern stable sets in matching problems. *Journal of Economic Theory*, 134(1):537–547, May 2007.

Thomas J. Miceli. The decision to regionalize in the provision of education: An application of the tiebout model. *Journal of Urban Economics*, 33(3):344–360, May 1993.

John Von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton university press, Princeton, 1944.

A. Pakes. Alternative models for moment inequalities. *Econometrica*, 78(6):1783–1822, 2010.

A Pakes, J Porter, K Ho, and J Ishii. Moment inequalities and their application. Technical report, 2011.

Debraj Ray and Rajiv Vohra. Equilibrium binding agreements,. *Journal of Economic Theory*, 73(1):30–78, March 1997.

Alvin Roth. Deferred acceptance algorithms: history, theory, practice, and open questions. *International Journal of Game Theory*, 36(3):537–569, March 2008.

Table 1: Summary Statistics (for municipalities, by type of prefecture)

	Units	Metropolitan	Mixed	Rural
$\tilde{c}$ (std. fiscal need)	¥100,000,000	361.86	64.42	53.85
		(776.37)	(182.15)	(169.60)
$\bar{\tau}Y$ (std. fiscal rev.)	¥100,000,000	281.60	40.84	24.75
		(665.12)	(146.98)	(112.47)
POPULATION	100,000	2.04	0.34	0.23
		(3.90)	(0.96)	(0.89)
AREA (sq. km.)	100	0.45	0.98	1.93
		(0.54)	(1.06)	(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
# municipalities		144	2486	625
# prefectures		3	36	8

Table 2: Dependent variable is v_{mS}, (structural) utility to muni m from merger S

	tax floor	no tax floor
CONSUMPTION (β_0)	200.99^{**}	98.91^{**}
	$(110.0, 440.0)$	$(60.0, 350.0)$
GOVERNMENT (β_1)	4.58^{**}	2.73
	$(2.0, 12.0)$	$(0.0, 9.0)$
DISTANCE (β_2)	-0.21^{**}	-0.25^{*}
	$(-0.35, -0.02)$	$(-0.40, 0.0)$
STONE_GEARY (β_3)	1.05^{**}	1.03^{**}
	$(0.95, 1.13)$	$(0.94, 1.25)$
EFF_OF_SCALE (β_4)	0.49^{**}	0.55^{**}
	$(0.38, 0.61)$	$(0.29, 0.68)$
N (prefectures)	47	47

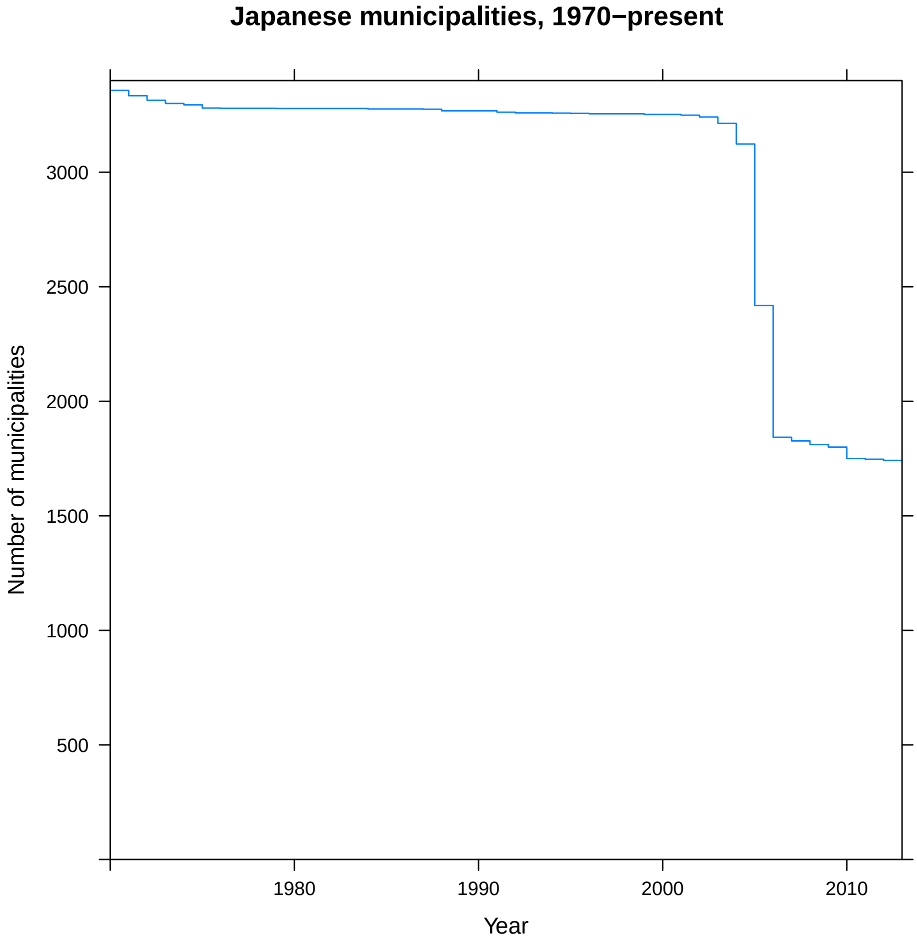
** 95% level

* 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

Figure 1:



Japanese municipalities, 1970−present
3000
2500
2000
1500
1000
500
Number of municipalities
1980
1990
2000
2010
Year

Figure 2: Prefectures of Japan

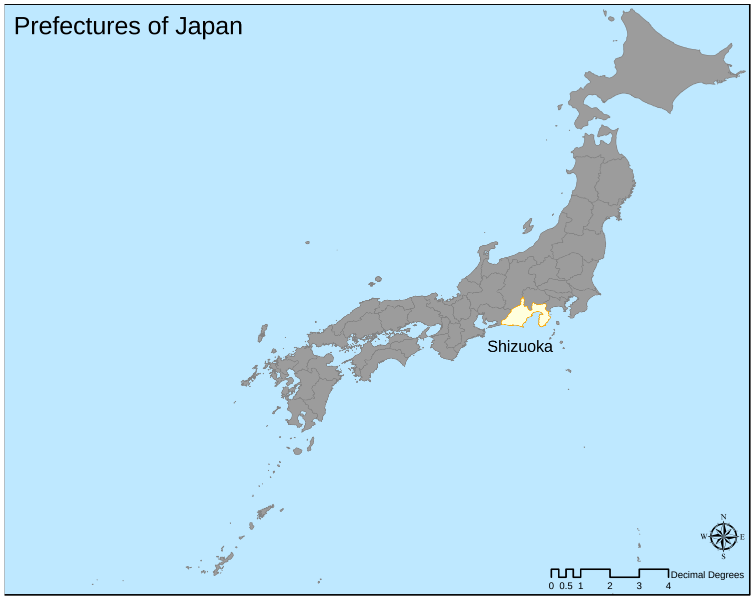


Figure 3: Shizuoka Prefecture

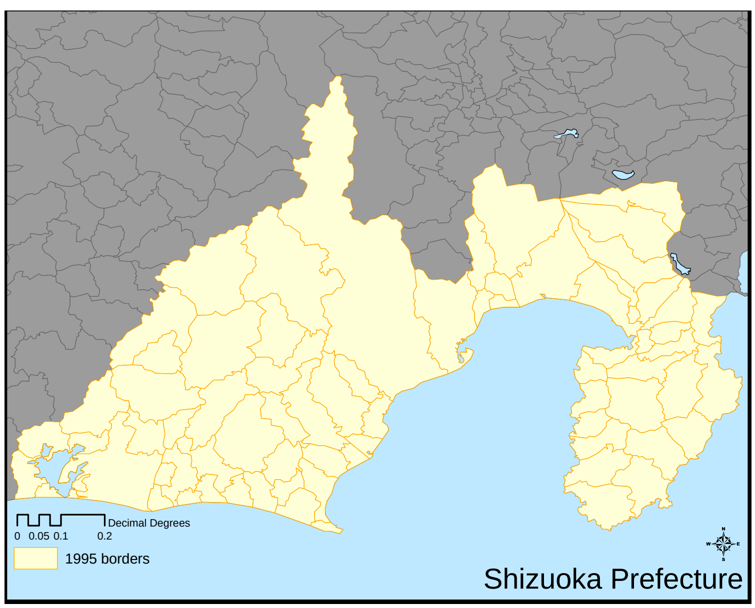
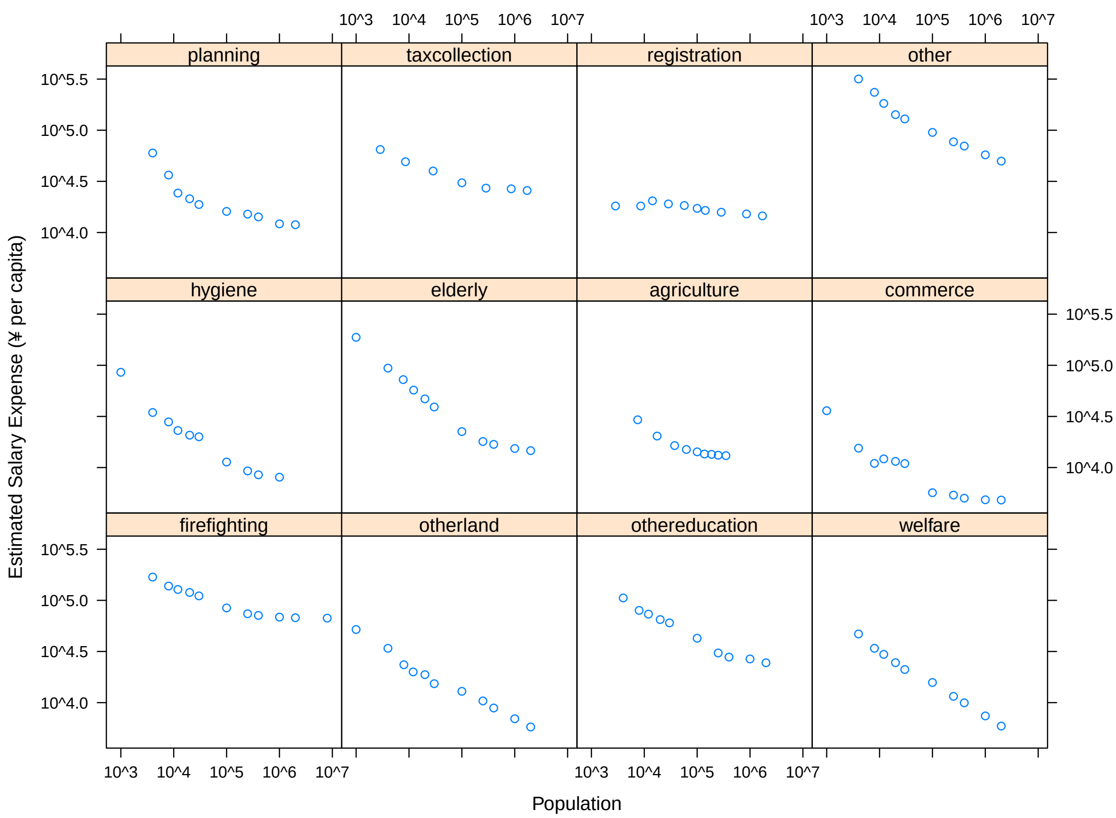
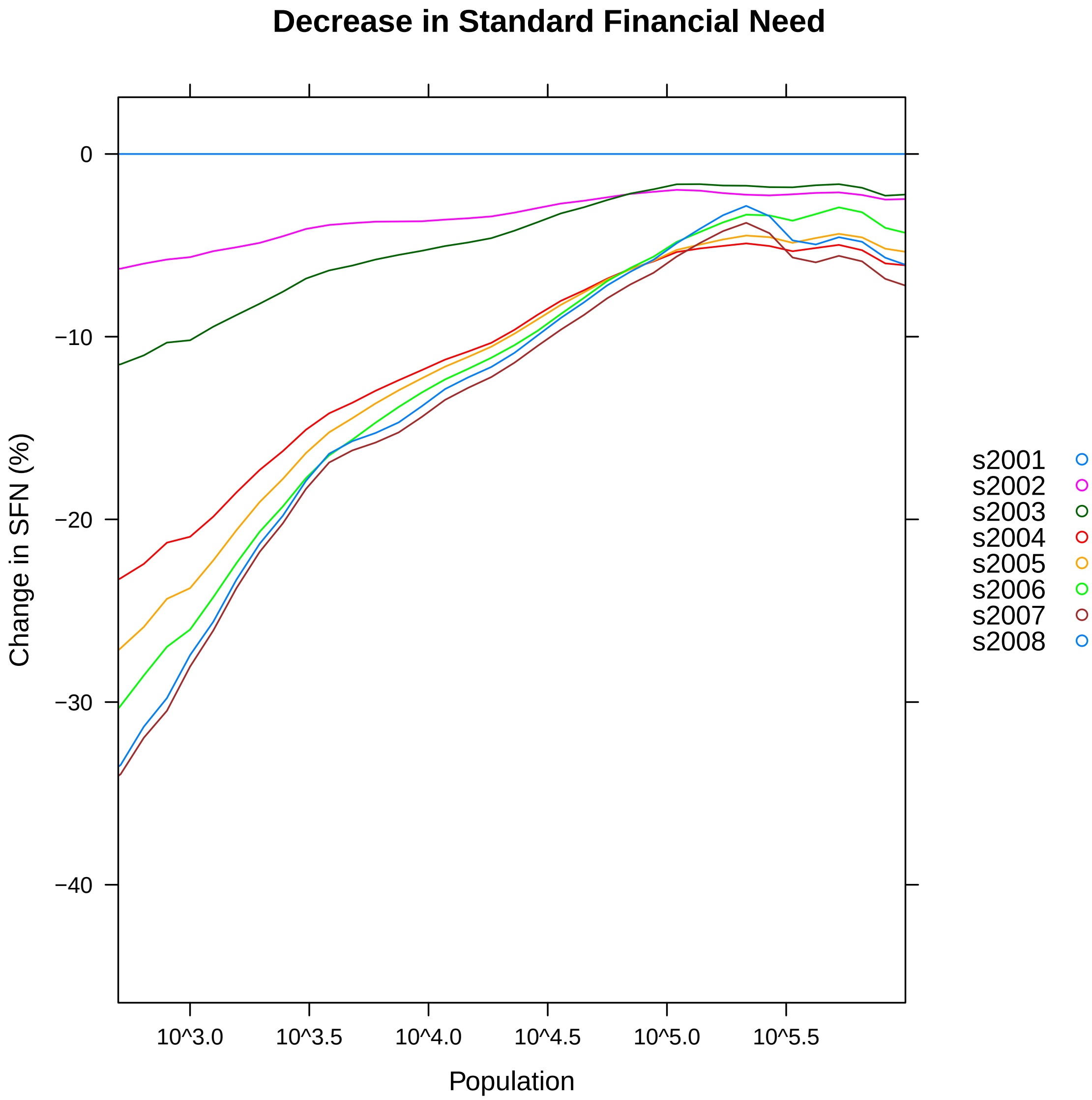


Figure 4: Adjustment coefficient based on number of personnel required



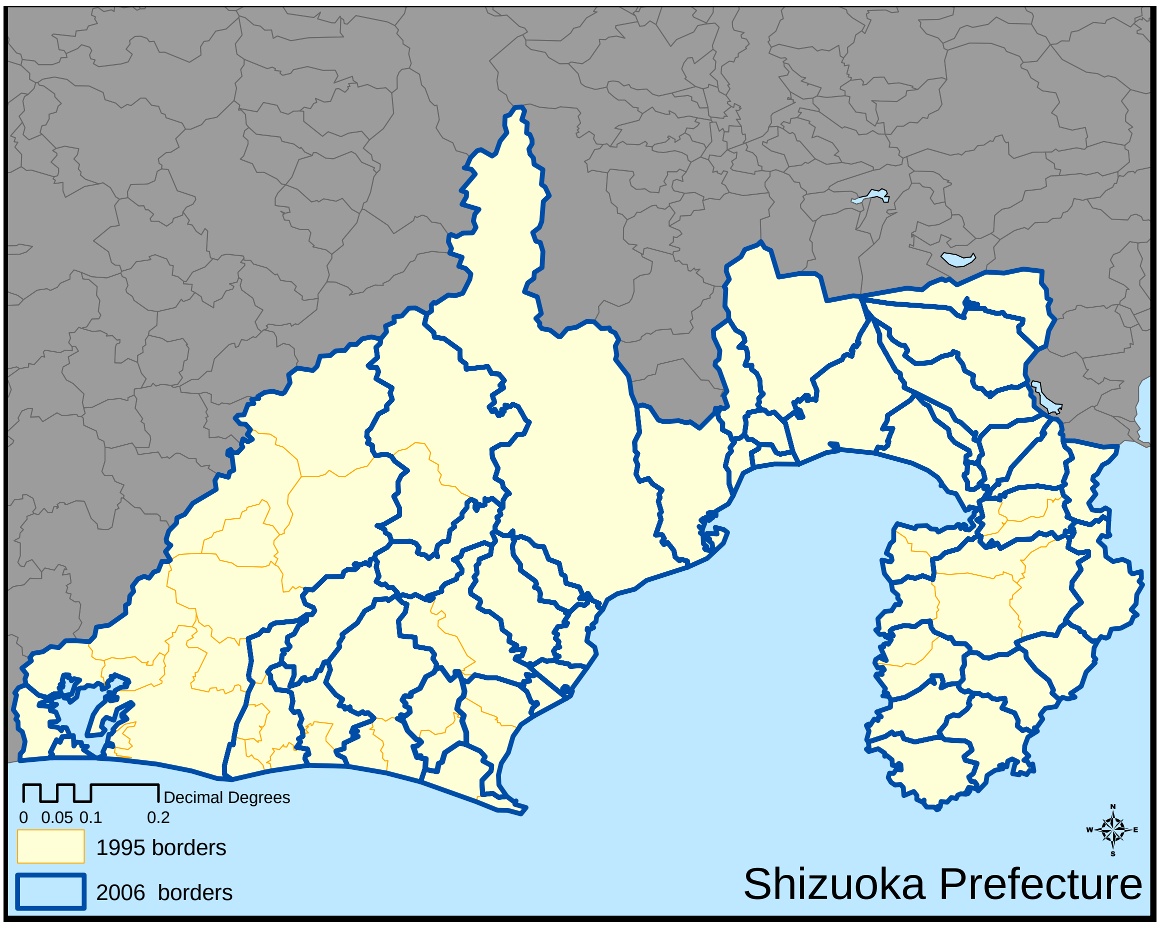
(Vertical axis is coefficient multiplied by estimated "standard" personnel expenditures)

Figure 5:



(Kernel regression using fixed municipality characteristics, and each year's transfer rule)

Figure 6: Mergers in Shizuoka Prefecture



A Voting Model

The variables determined by a political process at the local level, given a certain municipality m, are q_m, τ_m, and θ_m. The national government chooses the transfer function T. Due to the form of the utility function in Equation 3, all individuals agree on the optimal level for the public good, q_m^*, and the tax rate, τ_m^*. For the other political choices, however, different individuals will have different ideal points, and each of these may be multidimensional: the location θ_m of the public good is geographic (latitude and longitude), and the transfer function is chosen from a space of functions which in Section 4 will be assumed to be two dimensional. A very simple model of the political process at both the local and national level gives the result that the policy selected is a weighted sum of individual utilities. To obtain this result, make the following assumptions:

1. There are two identical office-motivated candidates that run on policy platforms that they can commit to.

2. Voting in elections is determined via a probabilistic voting model where vote probabilities are linear in utility difference between the two candidates.

3. There is a continuum of voters.

4. $\forall \theta, \theta', \forall \gamma \in (0, 1)$, the set of voters i such that

$$u_i(\gamma\theta + (1-\gamma)\theta') > \gamma u_i(\theta) + (1-\gamma)u_i(\theta')$$

   has positive measure.

With these assumptions, the unique political equilibrium is for both candidates to propose θ_m^* to maximize the sum of individual utilities: this is Theorem 4 in Banks and Duggan [2005].[43]

A.1 Japanese Context

Municipal politics in Japan involves both a mayor and a municipal council, and thus there is in reality more than the one decision maker assumed above; however, with the

[43] An additional requirement of Theorem 4 of Banks and Duggan [2005] is that for each voter, utility is (weakly) concave with respect to the policy choice. The utility function given in Equation 3 is concave in θ, although it is not strictly concave.

exception of very large "designated municipalities" the council is elected on an entirely at-large basis, without wards or other subdivisions. The mayor has veto power, which can be over-ruled by a 2/3rds vote of the municipal council. Given the lack of wards in the municipal council, it is not entirely clear how or why policies proposed by council might diverge from policies proposed by the mayor, although examples of this sort of conflict can be found in municipal records. Because this paper's focus is inter- rather than intra- municipal decision-making, the following assumption will be used: mayors will veto anything other than the policy proposed in their campaign, and less than 2/3rds of council will be opposed to this policy.[44]

National level politics are even more complex, and thus diverge even more from this simple model. The candidates in this case would be political parties, which commit to party platforms. Issues with single-member constituencies and multiple houses in the legislature, with different malapportionment in each house, are abstracted away from. Election of representatives is also abstracted away from: individuals in areas that are overrepresented are simply assumed to be able to cast more votes, and are thus weighted more heavily.

Assumptions 3 and 4 are not quite satisfied in the data actually used: there are a large but finite number of voters, and there are a few cases (generally in the smallest municipalities), where there are locations θ and θ' such that all voters are indifferent between randomization between the policies, versus a convex combination. The argument regarding Assumption 3 is simply that thousands of voters is "close enough" to a continuum. Regarding Assumption 4, violations of this assumption still result in candidates proposing policies that maximize social welfare, only these policies are no longer necessarily unique. For example, with the utility function in Equation 3 consider a municipality with exactly half of its population at one point, and exactly half at another point. Then any θ^* between these two points is welfare maximizing, and there is not a unique political equilibrium. Empirically, actual population distributions are never this evenly balanced, and a unique θ^* can always be computed. Distance enters the utility function in Equation 3 linearly, and

[44] Prior to the merger period, mayors were responsible for delivering hundreds of "agency delegated functions" from higher levels of government, making them bureaucrats as well as politicians, and making it possible (at least in theory) for central ministries to fire a mayor for not performing a delegated function according to specifications. "Agency delegated functions" were abolished during the merger period, and municipal policies are thus modeled as being determined by local residents through a political process.

thus these θ^* are points that minimize the sum of distances, points sometimes referred to as "generalized medians".

A.1.1 Municipal Mergers

The general political rules for municipal mergers were the following:[45]

0. Mayors of municipalities can create "voluntary merger committees" and "study committees" to gather information, but there are no regulations regarding these committees, and they are not necessary in order to proceed with a merger.

1. A petition for a specified merger from 2% of eligible voters (or the municipal council) in any single municipality forces an official response from all the municipalities included in the proposed merger, based on a debate in their municipal councils. Unanimous "yes" responses result in the creation of an "official merger committee". There is no requirement regarding previous voluntary committees or study committees.

2. If a municipal council rejects the proposed merger committee, a petition from 1/6th of eligible voters in the relevant municipality forces a referendum on the creation of the merger committee. A majority vote in the referendum overrides the council's rejection.

3. The merger committee produces reports on the financial situation of the municipalities and proposes some characteristics of the merger (eg. the name of the merged municipality). A majority vote in each municipal council is required to finalize the merger.[46]

The existence of an official referendum process during the planning stage but not at the final approval stage suggests that the best strategy for politicians opposed to a merger might have been to remain silent during the initial stage, but then prevent the final resolution from passing in council. Behaviour such as this did in fact occur in a small number of municipalities, but does not appear to have been particularly common or successful. First, the process of creating the merger committee generally attracted a considerable amount of attention, particularly in smaller municipalities. In cases where there was controversy, referendum turnout rates could exceed 90%. It

[45] This discussion ignores many details, such as the distinction between *hennyuu* municipalities, where bylaws are inherited from one of the merger participants (normally the largest city), and *shinsetsu* mergers, where bylaws and regulations are developed from scratch.

[46] In general, the division of a municipality was prohibited. In one case, such a split did occur, but both of the resulting municipalities were immediately merged with different neighbours.

was thus difficult for politicians facing a potentially controversial merger to prevent a referendum regarding the creation of the merger committee, and conditional on that referendum passing it was difficult to then vote against the final proposed merger. Furthermore, in cases where politicians did vote against mergers that appeared to have popular support, a hitherto seldom used recall process was employed to remove them from office via a majority vote in a recall referendum. Whereas there was only one recall referendum during the 1990s, there were at least 41 during the merger period.

A formal interpretation of these rules is somewhat difficult; however, a common element in all mergers is that they were approved by all municipalities in question, either via local referendum, or in the municipal council.[47] As council resolutions were subject to veto by the mayor, this paper will assume that the binding constraint on the behaviour of a municipality is the ability of its residents to recall the mayor. Suppose that there is perfect information regarding what mergers are feasible (i.e. will be approved by all other participants). The mayor proposes a merger for the municipality to participate in, or proposes remaining independent. A single challenger then appears, and similarly proposes a policy. With the probabilistic voting model presented above, the challenger will run on the same platform as the mayor, and the selected merger will maximize the sum of the utilities of residents.[48]

A potential objection here is that the costs of organizing a recall election could be large, and thus the mayor's incumbency advantage significant enough to allow merger proposals far from the optimal to be enacted.[49] There are two responses to

[47]In about a third of cases, referenda were held. Most of these were nominally consultative, but there is only one instance in which a municipal council voted opposite to a referendum result. This case was complicated due to multiple referenda with conflicting results as well as a number of of other procedural irregularities, and finally resulted in a recall of the mayor and a request to the prefectural governor to reverse the merger. The request for reversal was denied.

[48]The assumption here that mayors do not have a large effect on mergers might still seem suspicious. Kawaura [2010] investigates the effect of a mayor's length of tenure on merger configurations, and finds effects that are small and not statistically significant at the 95% level. While there is certainly anecdotal evidence that certain mayors may have obstructed certain mergers, there is no immediately obvious relationship in the aggregate data. The private incentive for municipal politicians to maintain the independence of their municipality in order to preserve their own employment is not as strong as might be anticipated. This is due to central government policies: for example, the length of service required to receive a pension were reduced for politicians in a municipality participating in a merger, and following the merger period the pension system was abolished, with a (disadvantageous) one-time payment to those who did not meet the standard 12 year length of service requirement.

[49]A recall referendum required a petition by between 1/6th and 1/3rd of residents.

this objection: first, the merger period was sufficiently long that at least one regularly scheduled election occurred during the merger period, and during this election the merger issue was particularly salient; second, the cost of organizing a recall does not appear to be as large as might be supposed. Specifically, in 4 of the 41 recalls, a majority voted against the recall in the referendum, and in another 6 of the recalls, the mayor was re-elected in the special election following the recall process (usually after resigning voluntarily to avoid the recall referendum). Thus, a full quarter of the organized recall referenda did not succeed in removing the mayor. If the costs of organizing a recall referendum were very high, one would expect that they would be organized only when the mayor would not have majority support in the recall referendum or the subsequent election. Thus, this paper will use the assumption that the municipal merger selected by each municipality maximized the utility of local residents, given the possible alternative mergers.[50]

B Stability Concept

Suppose that player $m \in M$ has preferences $\preceq_m$ defined over the set $\{S \subset M | m \in S\}$, 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 solution set is defined using the von Neumann and Morgenstern [1944] "stable set". Although the VNM stable set was originally defined in terms of imputations rather than coalition structures, this paper follows Lars [2007] 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 (Lars 2007)**.** *Let $<$ be a dominance operator, and $\Pi^{VNM} \subseteq \Pi$. Then Π^{VNM} is called a stable set for $(\Pi, <)$ if the following two properties hold:*

1. $\forall \pi, \pi' \in \Pi^{VNM}$, $\pi \not< \pi'$. *(Internal stability)*
2. $\forall \pi \notin \Pi^{VNM}, \exists \pi' \in \Pi^{VNM}$ *where* $\pi < \pi'$. *(External stability)*

[50]The specific definition of "alternative" that used here is that they would be deviations as discussed in Appendix B, below.

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.[51] The hedonic game considered in this paper is simpler than the "equilibrium coalition structures" that Ray and Vohra examine, and thus in this paper both refinements and coarsenings 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 $\pi \prec_S \pi'$, $S \in \pi'$, 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 favour of merging to create π', or there is a subset of "perpetrators" in π that are unanimously in favour of deviating from their current coalition. In the former case, π' is the coarsening that results from the merger, while in the latter 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. *$\pi \rightarrow \pi'$ if $\exists S$ such that either $\pi \nearrow_S \pi'$ or $\pi \searrow_S \pi'$, where*

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

The recursion is well defined since Q' is a proper subset of π'. Now let $\twoheadrightarrow$ be the transitive closure of $\rightarrow$.[52] Assume that $\Pi \neq \emptyset$.

Proposition 1. $\Pi^* = \{\pi | \nexists \pi'$ *such that* $\pi \rightarrow \pi'\}$ *is a stable set with respect to* $(\Pi, \twoheadrightarrow)$.

[51] An 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 multi-player 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.

[52] That is, $\pi \twoheadrightarrow \pi'$ if either $\pi \rightarrow \pi'$ or $\exists \{\pi_1, \ldots, \pi_n\}$ where $\pi \rightarrow \pi_1 \rightarrow \ldots \rightarrow \pi_n \rightarrow \pi'$. To see why the transitive closure is used here, consider the case where $\pi_1 \searrow_S \pi_2 \nearrow_{S'} \pi_3$. 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 $\pi_1 \not\rightarrow \pi_3$.

Proof. By construction, Π^* is internally stable. Now take some $\pi \notin \Pi^*$. Then $\exists\{\pi_1, \ldots, \pi_n\} \subset \Pi$ such that $\pi \to \pi_1 \to \cdots \to \pi_n$ 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:

(existence). Immediate by the above definition of Π^*. $\square$

(non-emptiness). 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. $\square$

(uniqueness). Suppose that Π^{**} is also a stable set with respect to $(\Pi, \twoheadrightarrow)$. Consider the bipartite directed graph defined by $\twoheadrightarrow$ with $\Pi^{**} \setminus \Pi^*$ and $\Pi^* \setminus \Pi^{**}$ as the two sets of nodes. Every node must have in-degree of at least one, but there can be no cycles. The only such graph is empty, and thus $\Pi^{**} = \Pi^*$.[53] $\square$

It can also be shown that Π^* contains a Pareto optimal partition:

(PO element). Let $\Pi^{\text{PO}} \subset \Pi$ be the set of Pareto optimal partitions, and $\rightsquigarrow$ the Pareto dominance operator. Suppose that $\Pi^{\text{PO}} \cap \Pi^* = \emptyset$ and consider the directed graph defined by $\twoheadrightarrow \cup \rightsquigarrow$ with Π^{PO} and Π^* as two sets of nodes. A cycle must exist, because $\forall \pi \in \Pi^{\text{PO}}, \exists \pi' \in \Pi^*$ such that $\pi \twoheadrightarrow \pi'$, but at the same time $\forall \pi \in \Pi^*, \exists \pi' \in \Pi^{\text{PO}}$ such that $\pi \rightsquigarrow \pi'$. Choose the starting point in this cycle such that $\pi_0 \rightsquigarrow \pi_1 \twoheadrightarrow \cdots \twoheadrightarrow \pi_n = \pi_0$. Let S_1^+ be the set of agents that strictly prefer π_1 to π_0. It cannot be that $\pi_1 \nearrow \pi_2$ because this is also a pareto improvement. Thus $\pi_1 \searrow_{S'} \pi_2$, and $S_2^+ = (S_1^+ \setminus R) \cup P$ where R is some subset of the residual, and $P \neq \emptyset$ is some subset of the perpetrators, and $(R \cup P) \subset S'$. Since $S_n^+ = \emptyset$, at some point the agents in S_2^+ must be made worse off. This can only happen via refinements, and only if there is a residual smaller than S_2^+. The latter, though, implies that either some subset of S_2^+ cannot be made worse off, or that S_3^+ will contain some new element. Thus, S^+ can never be empty. Thus a cycle cannot exist, and there is some Pareto optimal element in Π^*. $\square$

[53] To see this, attempt to iteratively construct a non-empty graph that has the desired form.

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.[54]

C Data and Institutional Details

Population data comes from the 1995 Japanese national census, which provides data at the kilometer grid square level. Information on municipal boundaries is taken from shape files produced by ESRI Japan, also for 1995. By combining these two data sources, the location of individuals in municipalities can be known to the kilometer grid square level.

To calculate distances, first, the population of grid squares that are on a boundary between two municipalities is divided between the municipalities in proportion to the area of each grid square in each municipality. Then, for any θ_m, the distance $\ell(i, \theta_m)$ can be calculated as the great-circle distance from the physical longitude-latitude location of individual i to θ_m. For computational simplicity, all individuals within a given census grid square are assumed to live at the centre of the square. The distances in question are small relative to the curvature of the earth, so this is effectively a straight-line distance calculation.

The location θ_m^* chosen by a municipality will minimize the sum of these individual distances, due to the assumption that the local political process is as described in Appendix A. These θ_m^* are calculated via standard optimization techniques. Although, as discussed in Appendix A, there are cases where θ_m^* might not be unique, a unique value is in fact obtained for all municipalities. For each coalition S, the optimal location θ_S^* is calculated via exactly the same process. The value of $\ell_m(\theta_S^*)$ in Equation 5 can then be calculated by averaging over distances $\ell(i, \theta_S^*)$ for all individuals in m. This process is computationally intensive, but $\ell_m(\theta_S^*)$ depends neither on ϵ nor on β, and thus the calculation of these distances only needs to be performed once.

Data for municipal characteristics X_m comes from the Statistical Information

[54] There may be some "solutions" that seem particularly unattractive: $\{\pi \in \Pi^* | \exists \pi' \in \Pi^*, \pi \rightsquigarrow \pi'\}$. 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.

Institute, which aggregates a variety of government sources. Where 1995 data was not available, the year closest to 1995 was used. Data regarding municipal mergers comes from the Japan Geographic Data Center. Figure 6 shows the mergers that occurred in Shizuoka Prefecture.

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 is not publicly available.[55] 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, nor by applying a specific set of *a priori* assumptions regarding efficiencies of scale.[56]

First, Ministry officials set $\bar{c}_k$ by considering the cost of providing component good k for a reference municipality: a city with population of 100,000, surface area of 160km^2, and other standard characteristics. The number and type of local bureaucrats necessary to provide the service is then estimated, along with the cost of equipment and materials, plus any transfers to the relevant target population (eg. child benefit payments). The number and type of bureaucrats that smaller and larger municipalities would require to provide the same level of service is then estimated.[57] National Personnel Authority salary scales are then 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.

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

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

[55]The 2009 version of the exposition of these formulae (the *Chihō Kōfuzei Seido Kaisetsu*), consists of 600 pages of Japanese legal text, 460 pages of formulae, and 240 pages of reference values.

[56]According to MIC officials, each year estimates are modified based on formal and informal feedback from municipalities and prefectures, observed spending patterns, and in-house research.

[57]The 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 4,000, 8,000, 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.

However, one pattern frequently observed is that $\widetilde{\widetilde{H}}_k$ takes the form

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

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.[58] 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 $X_{mk}\bar{c}_k$.[59] Thus, *de facto*, the method for calculating $\tilde{c}(X_m)$ is

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

where

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

which is Equation 18.[60] 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 total number of residents, but in some cases the relevant subgroup may be considered instead: the adjustment for services to the elderly is based on the number of residents over 65, the adjustment for agricultural services is based on the number of farmers, and so forth.[61] This adjustment is substantial, with the per capita cost of providing services usually estimated to be 2 to 3 times higher for a municipality of 4000 people than one of 100000. *Dankai* adjustments for some important services are shown in Figure 4. There are also other adjustments, such as one for population density: for example,

[58] The precise number of adjustment coefficients and component goods varies slightly from year to year.

[59] Expenses 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 is not available it is difficult to do this perfectly.

[60] X_m is more than a 24-tuple, with some elements used only in the calculation of the adjustment coefficients $\tilde{H}(X)$.

[61] The unit cost in these cases would be the estimated cost of providing elderly care for one person over 65, agricultural services for one farmer, and so forth.

the estimate for the cost of providing firefighting is increased from ¥1.009 million to ¥1.029 million if the population density of a city with population 100,000 were 150 per km^2 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 firefighting costs in general being greater than the effect on other component goods.[62]

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 is accomplished by modifying capital spending estimates, with the result that official municipal capital spending "needs" vary radically from year to year; estimates of the non-capital spending required to provide municipal services, on the other hand, change very little.[63] This sort of variation is captured in the model presented in Section 1 through 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 Construction Ministry, and involves pressure to promote spending on local projects. Over time, this has resulted in the addition of numerous "project" adjustment coefficients, each providing a special incentive for a variety of public works project. 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 3.

The tax base Y_m is determined from "Standard Fiscal Revenue" figures produced by the Ministry. While the model in Section 1 has each municipality choosing a single

[62]It 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 understates the Ministry's estimates of efficiencies of scale, making the finding that $\beta_4 < 1$ more surprising.

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

tax rate τ_m, actual municipal tax revenues come from several taxes, with "fixed asset" taxes (land, housing, and some business assets) and personal income tax being the two most important types. For each type of tax, the Ministry sets a "reference" tax rate, and then calculates the total revenue each municipality would receive if it charged these reference rates on its tax base. That is, if $\bar{\tau}^k$ is the reference rate for tax type k, and Y_m^k the tax base for this tax for municipality m, the Ministry "Standard Fiscal Revenue" estimate for municipality m is $\sum_k \bar{\tau}^k Y_m^k$. To convert this to the single tax base that is assumed in this paper, suppose that the single tax base is income, and set $\bar{\tau} = .12$, which is total municipal Standard Fiscal Revenue as a fraction of total income. Then calculate Y_m for each municipality so as to satisfy

$$\bar{\tau} Y_m = \sum_k \bar{\tau}^k Y_m^k. \tag{43}$$

That is, Y_m is calculated so that $\bar{\tau} Y_m$ is exactly equal to the Standard Fiscal Revenue for that municipality, as reported by the Ministry. The tax rate actually observed in the municipality, τ_m^*, is defined as taxes as a fraction of Y_m.

In general, collapsing multiple tax bases into a single tax base would be problematic, but in the Japanese case, although municipalities are *de jure* allowed to choose a tax rate different from the standard rate, the amount of actual variation is very low. For example, in the extreme case of Yuubari City, effectively bankrupt with a debt of over ¥3 million per capita, the income tax rate was raised from 6.0% to 6.5%, but almost all other municipalities charge the standard 6.0%. The standard fixed asset rate of 1.4% is levied by about nine out of ten municipalities, with the remaining tenth mostly charging 1.5% or 1.6%.[64] Thus, the observed tax data that the model is attempting to explain involves all municipalities charging very similar tax rates, equal or very close to the Ministry's reference rate $\bar{\tau}$. In particular, there are no cases where a municipality chose to charge a very high rate on a particular tax base for which the reference rate was much lower, a situation which could lead to high and nonsensical calculated values for τ_m^*.

[64] A fixed asset rate of more than 1.7% requires Ministry approval, but few municipalities are at this cap. While there are Ministry caps on taxes, these are rarely binding. The sole exception is for taxes on corporations, where a sizeable number of municipalities do charge at the upper bound. These corporate taxes are a small percentage of total taxes, and thus issues with this upper bound are not considered in this paper.

D Computational Details

D.1 First type of moment inequality

Choose the form of h in Equation 28 to be

$$h(\bar{\omega}_M, s_M|X) = \sum_{m,m' \text{ adjacent}} \frac{N_m N_{m'}}{\sqrt{N_m + N_{m'}}} (\bar{\omega}_m - \bar{\omega}_{m'})^2 + \sum_{m \in M} \frac{N_m - 1}{2} (s_m^2 - 1)^2 \quad (44)$$

where the first summation is over pairs of municipalities that are geographically adjacent.[65] The calculation of h^* in Equation 28 is 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 (29), a lower bound h^{**} can be used instead of calculating h^* directly.

Specifically, for an observed partition π, consider the following two types of deviations:

$$\mathscr{S}^{\mathrm{c}}(\pi) = \{S' | S' \text{ is a merger of two geographically adjacent singletons in } \pi\}$$

$$\mathscr{S}^{\mathrm{r}}(\pi) = \{S' | S' \subset S \in \pi\}$$

Here $\mathscr{S}^{\mathrm{c}}(\pi)$ are coarsenings of the partition π, and $\mathscr{S}^{\mathrm{r}}(\pi)$ are refinements. To retain computational feasibility, restrict $\mathscr{S}^{\mathrm{r}}$ to contain only coalitions that are geographically contiguous and do not cross more than two county boundaries.[66] Here, counties are defined using county definitions from the Meiji era.[67] For some larger observed

[65] This 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.

[66] For the purposes of determining geographic adjacency, islands with only a single municipality on them are treated as being connected to the closest municipality on the "mainland" (i.e. Hokkaidō, Honshū, Shikoku, or Kyūshū) if it is within 50km. Using this definition almost all observed mergers are geographically contiguous. More specifically, not counting island municipalities there are thirteen observed mergers that are not geographically contiguous. This is usually because one of the participants dropped out late in the merger process.

[67] In particular, the county boundaries used are from 1878 for eastern Japan, and 1896 for western Japan. Counties are statistical divisions, and have not had any political function since the 1920s. Counties in Tokyo and Nagano Prefectures are anomalously large, and thus in those prefectures only the restriction is to one and two counties, respectively, rather than three. Two actual mergers violate the restriction on number of counties that is imposed: one size twelve merger in Shizuoka, and one size eleven merger in Niigata. This represents 0.3% of all observed mergers.

coalitions S, the number of such subsets is still very large. In cases where there are more than 1000 potential subsets $S' \subset S$ for a given $S \in \pi$, use only the following:

1. S' is a singleton.
2. $S \setminus S'$ is a singleton (i.e. subsets that involve leaving out only one municipality).
3. A random sample of other subsets such that the number of subsets examined totals 1000.

The following are necessary conditions for π to be in the stable set $\Pi^*(\epsilon(\bar{\omega}_M, s_M)|X, \beta)$:

$$\begin{aligned} \forall S' \in \mathscr{S}^{\mathrm{c}}(\pi) \text{ either } & v_{mm}(\beta) + \epsilon_m(\bar{\omega}_M, s_M) > v_{mS'}(\beta) + \epsilon_{S'}(\bar{\omega}_M, s_M) \\ \text{or } & v_{m'm'}(\beta) + \epsilon_{m'}(\bar{\omega}_M, s_M) > v_{m'S'}(\beta) + \epsilon_{S'}(\bar{\omega}_M, s_M) \end{aligned} \tag{45}$$

where $S' = \{m, m'\}$ is the potential merger of two singletons, and

$$\forall S' \in \mathscr{S}^{\mathrm{r}}(\pi) \;\; \exists m \in S' \text{ s.t. } \; v_{mS}(\beta) + \epsilon_S(\bar{\omega}_M, s_M) > v_{mS'}(\beta) + \epsilon_{S'}(\bar{\omega}_M, s_M) \tag{46}$$

where S is the coalition that m is a member of in π. As $S' \subset S$, this condition can be expressed in a computationally simpler form:

$$\forall S' \in \mathscr{S}^{\mathrm{r}}(\pi) \;\; \max_{m \in S}(v_{mS}(\beta) - v_{mS'}(\beta)) > \epsilon_{S'}(\bar{\omega}_M, s_M) - \epsilon_S(\bar{\omega}_M, s_M) \tag{47}$$

Here $\mathscr{S}^{\mathrm{c}}$ considers only singletons in π, and $\mathscr{S}^{\mathrm{r}}$ considers only the non-singletons in π. Equations 26 and 27 show that the calculation of ϵ for the coalitions involved in these two sets of potential deviations depends on the $\bar{\omega}$ and s of disjoint sets of municipalities. This is because checking the necessary conditions for the coarsenings requires examining only those municipalities that remained as singletons, while checking the necessary conditions for refinements requires checking only those municipalities that participated in mergers. Furthermore, within the latter set, each actually observed merger can be checked separately. That is,

$$\begin{aligned} h^{**}(\pi|\beta, X) \quad = \quad & \min h(\bar{\omega}_M, s_M|X) \text{ s.t. (45) is satisfied} \\ & + \sum_S \min h(\bar{\omega}_M, s_M|X) \text{ s.t. (46) is satisfied for } S \end{aligned} \tag{48}$$

where conditions (45) and (46) depend on π, β, and X.

Next, note that for a given $\hat{\beta}$, the problem of calculating h^{**} can be represented as a minimization problem where the variables are $(\bar{\omega}_m - \bar{\omega}_{m'})^2$ for adjacent m and m', and s_m^2 for all m. The constraints in Equations 45 and 46 are linear in these variables, and the objective function h is a quadratic of them. Thus, the problem of computing h^{**} can be expressed as a quadratic program, or as a sum of the solutions to several quadratic programs as shown in Equation 48. This latter form is computationally feasible, and can be solved quickly using commercially available quadratic programming libraries such as CPLEX.

D.2 Second type of moment inequality

To see how Equation 32 can be used to create a moment inequality, first let $F_Q(\beta)$ be the distribution of the number of municipalities that would have participated in mergers during the 1999-2010 period if the government had not implemented any new merger promotion policies. By assumption (and after making an appropriate adjustment for the fact that the 1979-1998 period is longer than the 1999-2010 period) $F_Q(\beta)$ is stochastically dominated by $F_Q^{79}(\beta)$, the distribution of the number of municipalities that participated in mergers during the 1979-1999 period. F_Q is difficult to calculate directly: not only is the true equilibrium selection rule unknown but, as discussed at the beginning of Section 3.1, the precise membership of $\mathscr{S}$ is both unknown and likely very large. Thus, instead consider a bound μ_Q^*, such that

$$
\begin{aligned}
E_Q g_2(Q^{79}, \beta^0|X) &= \mu_Q^{79} - \mu_Q^*(\beta^0|X) \\
&\geq \mu_Q^{79} - \mu_Q
\end{aligned}
\tag{49}
$$

which is greater than zero because the expected number of mergers under no policy change has been assumed to not be greater that then number of mergers that occurred before the merger period.

This bound μ_Q^* can be computed by examining only size two mergers. Specifically, if $S = \{m, m'\}$, and $u_{mS} > u_{mm}$, $u_{m'S} > u_{m'm'}$, then any stable partition must have at least one of m and m' participating in a merger, because S is a blocking coalition for all other partitions. Let $\mathscr{S}_a$ be the set of size 2 mergers where the municipalities are geographically adjacent, and both municipalities prefer the merger to remaining as a singleton. This set can then be used to construct an easily computable minimal number of municipalities that must be involved in mergers. Consider the following variable, which is random because the membership of $\mathscr{S}_a$ depends on the draw of $\bar{\omega}_M$ and s_M:

$$Q^* = \underset{Q \subset M}{\operatorname{argmin}} \#Q \text{ s. t. } \forall S \in \mathscr{S}_a, S \cap Q \neq \emptyset \tag{50}$$

That is, Q^* is a minimal hitting set for $\mathscr{S}_a$: for each potential geographically contiguous size 2 merger where both participants prefer the merger relative to not merging at all, at least one of those municipalities is in Q^*. For a given random draw of shocks, Q^* can be computed via a linear program. Let $F_Q^*(\beta)$ be the distribution of $\#Q^*$. F_Q stochastically dominates F_Q^*, because $\mathscr{S}_a$ is a subset of all mergers whose participants prefer the merger to remaining as a singleton, and thus any stable partition must include at least $\#Q^*$ municipalities participating in mergers regardless of equilibrium selection rule. Examining potential pairwise mergers thus gives a lower bound on the total number of municipalities participating in mergers of any kind.

The calculation of Q^* involves a high dimensional integral because there are two idiosyncratic shocks for each municipality, and many municipalities in a prefecture. However, as only size 2 mergers between geographically adjacent municipalities are considered, and very few of these mergers actually occur, the interactions between municipalities that are far away are very weak. Thus, in a large prefecture such as Hokkaidō, with 210 municipalities, simulation error in the southern portion of the prefecture will have little effect on mergers of northern municipalities, and thus the law of large numbers will apply. In addition, there are multiple prefectures, and there are assumed to be no interactions between them.

D.3 Identified Set and Confidence Sets

The data consists of 47 prefectures, which are treated as independent coalition formation games. As in Table 1, prefectures are classified as "metropolitan", "mixed", and "rural" depending on the percentage of municipalities with a population of less than 10000. Let these sets of prefectures be J^{metro}, J^{mixed}, and J^{rural}, respectively. Let $\bar{g}_1^{\text{metro}}(\beta)$ be the sample moment of g_1 with prefectures J^{metro}:

$$\bar{g}_1^{\text{metro}}(\beta) = \frac{1}{\#J^{\text{metro}}} \sum_{j \in J^{\text{metro}}} g_1(\pi_j^0, \beta | X) \tag{51}$$

where π_j^0 is the actually observed partition in prefecture j.[68] Construct $\bar{g}_2$ similarly: for example, $\bar{g}_2^{\text{mixed}}$ would be

$$\bar{g}_2^{\text{mixed}}(\beta) = \frac{1}{\#J^{\text{mixed}}} \sum_{j \in J^{\text{mixed}}} g_2(Q_j^{79}, \beta | X) \tag{52}$$

where Q_j^{79} is the actually observed number of municipalities participating in mergers in prefecture j during the 1979-1998 period. There are thus three sample moments calculated from g_1 (metro, mixed, and rural), and another three in the same way for g_2. On the other hand, g_3 is only calculated for "metro" (it will not bind for any value of β for mixed or rural, due to the large number of mergers in these types of prefectures):

$$\bar{g}_3^{\text{metro}}(\beta) = \frac{1}{\#J^{\text{metro}}} \sum_{j \in J^{\text{metro}}} g_3(Q_j^{99}, \beta | X) \tag{53}$$

where Q_j^{99} is the actually observed number of municipalities participating in mergers in prefecture j during the 1999-2010 period. These sample moments are then combined into the test statistic

$$\begin{aligned} \mathbf{T}(\beta) = [\bar{g}_1^{\text{metro}}(\beta)]_-^2 + [\bar{g}_1^{\text{mixed}}(\beta)]_-^2 + [\bar{g}_1^{\text{rural}}(\beta)]_-^2 + \\ [\bar{g}_2^{\text{metro}}(\beta)]_-^2 + [\bar{g}_2^{\text{mixed}}(\beta)]_-^2 + [\bar{g}_2^{\text{rural}}(\beta)]_-^2 + \\ [\bar{g}_3^{\text{metro}}(\beta)]_-^2 + [\bar{g}_4(\beta)]_-^2 \end{aligned} \tag{54}$$

where $[x]_- = \min(x, 0)$.[69] Due to the distributional assumption already made regarding $\bar{\omega}$ and s, the first six of the these terms are uncorrelated. $\bar{g}_3^{\text{metro}}$, however, could well be correlated with $\bar{g}_1^{\text{metro}}$, as π^0 affects both of these sample moments. Similarly, $\bar{g}_4$ could plausibly be correlated with other moments. These correlations are estimated using the subsample approach given in Andrews and Guggenberger [2009], with critical values of $\mathbf{T}(\beta)$ computed accordingly. The identified set is the set of $\hat{\beta}$

[68] In the actual computations the weights on individual prefectures differ to account for the differing number of municipalities in each prefecture.

[69] Here g_4 actually consists of two moment inequalities, because both an intercept and the size of municipalities in question is used with γ in Equation 35, and both of these should have a coefficient of zero. Also, an adjustment is made to g_2 and g_3 to take into account that mergers are infrequent, and thus are not well approximated by a normal distribution: quantiles of the actual distribution of the lower bound on the number of mergers are estimated, and these are then used to perform a normalization.

such that

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}}\, \mathbf{T}(\beta). \tag{55}$$

Optimization is performed via Nelder-Mead, using the base implementation in R. For the second and third types of moment inequalities, 1000 simulation draws are used to calculate the moment. The optimization algorithm used finds only a local minimum, and thus there is the possibility that this differs from the global minimum. The calculation of the confidence set for each parameter, however, involves rerunning the optimization routine with the relevant parameter fixed at a certain value. On average about 100 values were used: some of these values were very close to the estimated parameter, while others were further away. These additional starting points do not result in any lower objective function values than the one that I find at the reported optimum.

Following convention, the 95% confidence set will be

$$\{\beta | \mathbf{T}(\beta) < \mathbf{T}(\hat{\beta}) + \mathbf{T}_{0.95}(\hat{\beta})\} \tag{56}$$

where $\mathbf{T}_{0.95}(\hat{\beta})$ is the 0.95 quantile of the distribution of the test statistic under the hypothesis that $\beta = \hat{\beta}$. Constructing a 95% confidence set for β is simplified because assumptions regarding the distribution and most of the correlation structure of the error terms have already been necessary in order to develop the model.

The coverage probability of this confidence set is checked via Monte Carlo. 100 simulations of the municipal merger process are performed using the estimated parameters as the true parameters. The 95% confidence set has a coverage probability of 1.00, and an 80% confidence set has a coverage probability of 0.98.[70]

D.4 Example with two municipalities per prefecture

The moment inequality based on g_1 has not previously appeared in the literature. A simple example illustrates the identified set produced: notably, it will not exclude $\beta = 0$.

Let J be a set of prefectures each containing only two municipalities, A_j and B_j, with a potential merger $S_j = \{A_j, B_j\}$. For simplicity, let there be only a single

[70] In these simulations, an average of 67 municipalities participated in mergers under the original government transfer policies. This corresponds to about 2% of the 3255 original municipalities.

idiosyncratic shock ϵ_j in each prefecture:

$$u_{A_jA_j} = u_{B_jB_j} = \beta \tag{57}$$
$$u_{A_jS_j} = u_{B_jS_j} = 2\beta + \epsilon_j \tag{58}$$

and make the distributional assumption $\epsilon_j \sim N(0,1)$. Here, setting the variance to one normalizes the scale for β. Let π^0 be the observed partition: the only options for prefecture j are the singletons $\{A_j\}$ and $\{B_j\}$, or the merger $\{A_j, B_j\}$. Define the following stability restriction:

$$R(\epsilon, \pi|\beta): \quad \forall j \in J, \quad \epsilon_j \leq -\beta \text{ if } \{A_j\}, \{B_j\} \in \pi \tag{59}$$
$$\epsilon_j \geq -\beta \text{ if } \{A_j, B_j\} \in \pi \tag{60}$$

Now choose $h(\epsilon) = \sum_{j \in J} \epsilon_j^2$. Thus, $E(h) = J$. Then define

$$\epsilon^*(\beta) = \operatorname*{argmin}_{\epsilon} h(\epsilon) \quad \text{ s.t. } R(\epsilon, \pi^0|\beta). \tag{61}$$

That is, $\epsilon^*(\beta)$ is the vector of idiosyncratic shocks that generate the least extreme value from h while still rationalizing π^0. Let that value of h be

$$h^*(\beta) = \min_{\epsilon} h(\epsilon) \quad \text{ s.t. } R(\epsilon, \pi^0|\beta). \tag{62}$$

Let ϵ^0 be the actual epsilons that were drawn and resulted in π^0 being observed. Let β^0 be the true value of β. Then

$$h^*(\beta^0) \leq h(\epsilon^0) \tag{63}$$

because ϵ^0 being drawn resulted in partition π^0 occurring, and thus $R(\epsilon^0, \pi^0|\beta^0)$ must be satisfied. If the inequality is always satisfied, it is satisfied in expectation:

$$E(h(\epsilon)) - E(h^*(\beta^0)) \geq 0, \tag{64}$$

where $E(h^*(\beta^0))$ indicates the expected value of h^* for a partition generated from a random draw of ϵ. In this particular example, for any draw of ϵ there will only be one stable partition, but neither this uniqueness nor any particular assumptions regarding an equilibrium selection rule is required for the above inequalities to hold.

The expected fraction of prefectures with a merger is $\Phi(\beta^0)$. The expected value of $h^*(\beta)$ is

$$\Phi(\beta^0)\min(0,\beta)^2 + (1-\Phi(\beta^0))\max(0,\beta)^2 \tag{65}$$

Using only the above moment inequality for identification, we will have the identified set

$$\{\beta \quad | \quad 1-\Phi(\beta^0)\min(0,\beta)^2 + (1-\Phi(\beta^0))\max(0,\beta)^2 \geq 0\} \tag{66}$$

which corresponds to the interval $[\frac{-1}{\sqrt{\Phi(\beta^0)}}, \ \frac{1}{\sqrt{1-\Phi(\beta^0)}}]$. This interval contains zero, which is a general property of this type of moment: the "entirely idiosyncratic" null hypothesis will never be rejected. In order to reject $\hat{\beta}=0$, some additional moments of some other type must be used. In the estimator in the main paper, these correspond to the moments comparing the expected number of mergers if the government had not changed any transfer policies to the actual number of mergers observed during the period in which the old transfer policies were in effect. Considering the specification used in the main analysis, at $\hat{\beta}=0$, there would have been a large number of "random" mergers, and thus this null hypothesis is easy to reject using these other moments.[71]

[71] The above estimator may appear to be somewhat similar to other estimation approaches, such as maximizing the probability that the observed partition is stable. Estimators in this latter set, however, are not in general consistent, and will thus not necessarily be inside the identified set based on the moment inequality used above. To take an extreme example, suppose that there are K municipalities in each prefecture, and that the only mergers that are possible are $\{m_1, m_2\}$, $\{m_1, m_2, m_3\}$, ... $\{m_1, m_2, ..., m_K\}$. Preferences are determined by

$$u_{m_k S} = \frac{\#S}{K} + \sqrt{K}\epsilon_j, \quad S \neq \{m_k\} \tag{67}$$

$$u_{m_k m_k} = k - \frac{K-1}{2} \tag{68}$$

where $\#S$ is the number of municipalities in S, and j indexes prefectures. Thus there is again only one idiosyncratic shock per prefecture, and only one stable partition: if $\epsilon = 0$, for example, municipalities up to K/2 will merge. The probability of any given merger being stable is thus small if K is large. Now consider the more general model

$$u_{m_k S} = (1-\beta)(\frac{\#S}{K} + K^{1/3}\epsilon_j) + \beta(k - \frac{K+10}{2} + \epsilon_j^2), \quad S \neq \{m_k\} \tag{69}$$

$$u_{m_k m_k} = k - \frac{K-1}{2} \tag{70}$$

When $\beta = 1$, then, most of the time no mergers are stable, but when there is a very extreme ϵ then all mergers are stable (indifferent municipalities prevent deviations in the definition of stability used in this paper). Now suppose that the true value of β is $\beta^0 = 0$, and consider a pseudo-likelihood estimator that maximizes the probability that the observed partition is stable. Then for sufficiently large K, $\beta^{\text{PMLE}} = 1$, because all mergers are stable when any merger is stable, unlike the situation at β^0. The identified set using the technique outlined above, however, would exclude $\hat{\beta} = 1$, as this would involve ϵ^* that are very extreme, and the moment inequality would be violated. Thus, the (inconsistent) maximizer of the pseudo-likelihood function based only on stability is not always contained in the identified set.