ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208629
New Haven, CT 06520-8269
http://www.econ.yale.edu/~egcenter/

CENTER DISCUSSION PAPER NO. 997

Political Mergers as Coalition Formation

Eric Weese
Yale University

April 2011

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

This paper originated as a thesis chapter. I would like to thank my thesis committee: Daron Acemoglu, Abhijit Banerjee, and Esther Duflo. I would also like to thank attendees at seminar presentations for their helpful comments, and the members of the MIT Political Economy working group for support, encouragement, and further useful suggestions. This research was supported by a Canadian Institute for Advanced Research junior fellowship and a Japan Society for the Promotion of Science postdoctoral fellowship. The latter was held at Hitotsubashi University on the invitation of Masayoshi Hayashi, who, along with Masashi Nishikawa, explained to me the details of the Japanese local public finance system. Thanks to Vadim Marmer and Francesco Trebbi for pointing out an econometric error in an earlier version of this paper, and Alberto Alesina and Nancy Qian for suggestions regarding the structure and text. Konrad Menzel offered valuable comments regarding econometric strategies. Computational support was provided by the Yale Faculty of Arts and Science High Performance Computing facilities, based on initial computations performed at MIT.

Political Mergers as Coalition Formation*

Eric Weese†

March 2011

Abstract

Political coalition formation games can describe the formation and dissolution of nations, as well as the creation of coalition governments, the establishment of political parties, and other similar phenomena. These games have been studied from a theoretical perspective, but the resulting models have not been used extensively in empirical work. This paper presents a method of estimating political coalition formation models with many-player coalitions, and then illustrates this method by estimating structural coefficients that describe the behaviour of municipalities during a recent set of municipal mergers in Japan. The method enables counterfactual analysis, which in the Japanese case shows that the national government could increase welfare via a counter-intuitive policy involving transfers to richer municipalities conditional on their participation in a merger.

JEL codes: C63, D71, H77

Keywords: Computational Techniques, Coalitions, Municipalities

*This paper originated as a thesis chapter. I would like to thank my thesis committee: Daron Acemoglu, Abhijit Banerjee, and Esther Duflo. I would also like to thank attendees at seminar presentations for their helpful comments, and the members of the MIT Political Economy working group for support, encouragement, and further useful suggestions. This research was supported by a Canadian Institute for Advanced Research junior fellowship and a Japan Society for the Promotion of Science postdoctoral fellowship. The latter was held at Hitotsubashi University on the invitation of Masayoshi Hayashi, who, along with Masashi Nishikawa, explained to me the details of the Japanese local public finance system. Thanks to Vadim Marmer and Francesco Trebbi for pointing out an econometric error in an earlier version of this paper, and Alberto Alesina and Nancy Qian for suggestions regarding the structure and text. Konrad Menzel offered valuable comments regarding econometric strategies. Computational support was provided by the Yale Faculty of Arts and Science High Performance Computing facilities, based on initial computations performed at MIT.

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

1 Introduction

Issues related to political coalition formation have recently attracted considerable interest from theorists and policy makers. For example, the formation and dissolution of countries can be seen as a political coalition formation game, with coalitions consisting of residents of a geographic area [Alesina and Spolaore, 1997]. The question of how many countries ought to exist and where borders should be drawn has current relevance in places such as Georgia or Sudan, and recent episodes such as the dissolution of Yugoslavia can be explained using a coalition formation framework [Desmet, Breton, Ortuno-Ortin, and Weber, 2009]. The formation of a government also corresponds to a political coalition formation game, with the political parties being the players. Similarly, parties could themselves be viewed as resulting from an underlying political coalition formation game, this time with individual legislators as the players [Merlo, 2006].

Changes to the rules governing the coalition formation game may lead to a different and more efficient coalition structure. Any analysis of the effect of such changes requires knowledge of the underlying structural parameters and an understanding of the process of coalition formation given various possible sets of rules.[1] In order to evaluate different potential rules it is first necessary to develop a model of the behaviour of the players participating in the coalition formation game. This model can then be used to predict the changes in behaviour that would result from the imposition of different sets of rules.

Although models of coalition formation date back at least to von Neumann and Morgenstern [1944], relatively few empirical papers have made use of such models, and these papers have tended not to examine the effect of possible changes in the rules of the coalition formation game being studied.[2] Desirable properties of some specific forms of coalition formation games, such as two-sided matching games, have led to extensive empirical study of those game forms [Roth, 2008], but empirical research on more general coalition formation models has been hampered by the fact

[1]For example, a recurring proposal in Canada is that members of parliament should be required to stand for a by-election if changing their party affiliation between general elections. Had this rule been in place during recent parliaments, different coalition structures might have resulted, leading to different governments and policy outcomes. Similarly, different laws regarding how municipalities can cooperate to provide public goods, or how farmers can establish agricultural cooperatives, could lead to different coalition structures with different welfare implications.

[2]Diermeier, Eraslan, and Merlo [2003] is an interesting exception.

that neither existence nor uniqueness of a stable coalition structure is guaranteed, and the number of coalition structures increases exponentially with the number of players. Recently, Fox [2008] and related papers have studied coalition formation games with transferrable utility, which is particularly relevant for issues in industrial organization. This paper studies games where transfers are not possible, the case which is more relevant for issues in political economy [Acemoglu, 2003].

This paper presents a method of estimating the structural parameters of a political coalition formation model and then applies it to a recent set of Japanese municipal mergers. In the *Heisei Daigappei*, individual Japanese municipalities could choose what merger if any they wished to participate in, given a fixed set of national government transfer policies. The parameters that determine municipal preferences over mergers are estimated, and these estimates are then used to predict the effect of alternative national government transfer policies. The *Heisei* mergers are particularly attractive from a modelling perspective, as government policy led to mergers occurring only during 1999-2010, and thus the resulting coalition structure can plausibly be treated as the outcome of a single period coalition formation game.[3] Furthermore, the mergers are of interest from a policy perspective, since due to efficiencies of scale the smaller municipalities spend over ¥1,000,000 per capita per year providing the same services that larger municipalities provide for slightly over ¥100,000, and almost all of this difference was being subsidized by the national government.[4] Thus, in addition to a new method of analyzing political coalition formation games, the paper also contributes the specific analysis of an economically interesting case.

The methodological contribution consists of the use of simulated maximum likeli-

[3]In general, a problem with applying political coalition formation models to observed data is that political coalitions once formed tend to persist, and changes that do occur are often separated by large time periods. The extremely high cost of any realignment means that the stability of existing borders does not provide much information, and it is not clear what it means for there to be a "stable" coalition structure, if changes to this structure occur over time at a slow but constant rate. The Japanese data used in this paper mostly avoids this problem. The fiscal crisis of the 1990s precipitated such significant changes in intergovernmental transfers that in many cases the old municipal borders were effectively untenable, thus leading to a very large number of mergers during the window when mergers were allowed. Furthermore, during the 1970-1995 period, national policy had made municipal mergers extremely unattractive, and thus boundaries remained effectively unchanged even though demographic changes were rendering these boundaries increasingly inefficient.

[4]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.

hood estimation to obtain structural parameters describing players' preferences over coalitions when the observed coalition structure can be treated as the outcome of a Bogomolnaia and Jackson [2002] hedonic coalition formation game. Two ways of overcoming problems related to non-existence or multiplicity of stable coalition structures are considered. The first method is to assume that all players have the same preferences over coalitions, resulting in the existence of a unique stable coalition structure [Farrell and Scotchmer, 1988]. This is the approach used by Gordon and Knight [2009], but replacing their method of moments estimator with a simulated maximum likelihood estimator makes it computationally feasible to examine mergers involving more than two players.[5] However, the restriction that all players must have the same preferences over coalitions does not work well in the Japanese case, where many local government services are provided at a single specific physical location (city hall, library, health centre, etc.). Surveys suggest that residents of municipalities at the geographic "edge" of a proposed merger were concerned about the distance to post-merger public facilities, while residents of more centrally located municipalities were not. This implies an Alesina and Spolaore [1997] style model, where players' ideal points are distributed over a geographic policy space. The arrangement of jurisdictions is then the result of a tradeoff between economies of scale in the provision of public goods, and heterogeneity in preferences over the location of those goods.

A second strategy is thus developed, which allows players' preferences over coalitions to differ, but restricts the types of blocking coalitions that can form. This guarantees existence but not uniqueness of a stable coalition structure [Ray and Vohra, 1997], and thus estimation requires an assumption regarding which one of the set of stable coalition structures is actually selected. The advantage of this approach, however, is that the utility function can include interactions between individual and coalition characteristics, and thus geographic distance to a Banks and Duggan [2005] generalized median voter can be included.[6] This provides a direct link between a

[5]This is because the method of moments approach requires computing the stable coalition structure repeatedly as part of the estimation process. The computational advantage of the simulated maximum likelihood approach, however, does not necessarily hold for all datasets: in some cases a very large number of simulations might be required to reduce simulation bias to an acceptable level, thereby making the method of moments approach more attractive. In the case of the Japanese data used in this paper, SML estimates even with very few simulations (eg. $R = 10$) appear to give acceptable results.

[6]The distributional assumption regarding idiosyncratic preferences also differs, with shocks being i.i.d by municipality by coalition, rather than i.i.d by coalition as in the first approach. The correlation between idiosyncratic preferences regarding a given coalition is thus 0 rather than 1; however,

theoretically consistent model of jurisdiction formation and the estimating equation as actually implemented.

These two strategies are then applied to Japanese municipal merger data, with the cost of providing public services derived from existing national government estimates. Geographical features of the data allow the set of possible coalitions to be reduced to the point where the model is computationally tractable. The resulting parameter estimates differ depending on the estimation strategy used. Some of the coefficients estimated by restricting preferences (following Farrell and Scotchmer [1988]) have opposite sign to what would be expected, while those estimated by relaxing the requirements for stability (following Ray and Vohra [1997]) have the expected sign and plausible magnitudes. Monte Carlo exercises show that a majority of this difference can be explained by the inability to include relevant variables within the restricted preferences framework.[7]

In general, in non-transferable utility coalition formation games there are often coalitions that, if formed, would increase the utility of some players by large amounts, but these coalitions do not form because some other participants in the coalition would end up with slightly lower utility. Thus, national government intervention could lead to different and better coalition structures forming. The structural parameters estimated are used to examine the effects of two counterfactual policies. First, the possibility of national government enforcement of transfers is considered, where the national government allows decentralized negotiations over these transfers to take place between municipalities. In this case, where the game is converted into a transferable-utility game, the outcome depends on the bargaining power of different types of municipalities. While this policy increases the number mergers that occur, it also leads to potentially very large transfers from poor municipalities to richer ones, and the exact amount of the transfers cannot be known in advance without knowing the bargaining method by which municipalities divide the benefits of a merger. Even under the most optimistic assumptions regarding bargaining power, the poorest municipalities end up worse off than under the original policy.

it is not clear that this is an improvement.

[7] An earlier version of this paper [Weese, 2008] focused on variables that Kido and Nakamura [2008] and other previous studies had identified as important: population, surface area, etc.. Plausible coefficient estimates for those variables were obtained using the restricted preferences approach, but it is very difficult to respond to the Lucas critique while still satisfying the restricted preferences conditions.

Next, an alternative is considered where the national government provides an additional financial incentive for municipalities to participate in mergers. A sample budget-balanced policy results in higher utility (equivalent on average to an increase in income of 0.4%) for both poor and rich municipalities, even though incentives to participate in mergers are only offered to relatively richer municipalities. This result is somewhat counter-intuitive as the problem the national government was attempting to solve was the high cost of supporting small, poor municipalities. A regressive conditional transfer – taxing everyone and transferring money to the residents of richer municipalities that participate in mergers – is not an obvious response to this problem. The result is consistent with theory, however, since an incentive for richer municipalities to merge with their neighbours will lead to those neighbours benefitting from higher levels of public goods. Providing an incentive to richer municipalities mimics the transfers that the municipalities themselves offered in the transferable utility game, but with amounts that are not as large. Thus, fewer mergers occur, but the poorer municipalities are on average better off than in the transferable utility case because they do not have to pay huge transfers to richer municipalities.[8]

The major contribution of this paper is to develop an empirical framework for the estimation of political coalition formation models that takes into account theoretical characteristics of solutions and allows for the analysis of counterfactual policies. In addition to Gordon and Knight [2009], closely related papers are Brasington [1999] and Saarimaa and Tukiainen [2010], which use the maximum likelihood estimator from Poirier [1980] and are thus restricted to considering each pairwise merger in isolation from other potential mergers.[9] Other recent empirical political coalition formation papers generally focus on describing patterns that are observed in political boundaries, while this paper estimates structural parameters and predicts how counterfactual policies would change the set of boundaries forming.[10] With suit-

[8]This result is related to the theory presented by Armstrong and Vickers [2010] regarding anti-trust regulation of corporate mergers; however, in the Armstrong and Vickers model, the cost of allowing certain mergers to happen is that other, better, mergers do not occur, whereas in the model presented below, the primary cost of having more mergers occur is the ever larger transfers to richer municipalities that must be provided. The transferable utility case thus has "too many" mergers, at least for some social welfare functions.

[9]In models of the type used by Brasington and others, the probability that players 1 and 2 will form a coalition is unaffected by the other options that 1 or 2 might have. The method presented below and that used by Gordon and Knight appear to be the only ones that take into account that the presence of a player 3 and an attractive $\{1, 3\}$ coalition may disrupt a $\{1, 2\}$ coalition that would otherwise form.

[10]Most empirical studies of political mergers thus far focus on American school districts. Miceli

able modifications, the method used in this paper could be applied to other types of coalition formation games, possibly in other fields as well as in political economy.

The rest of the paper has the following structure. The general estimation strategy is presented in Section 2, including both the version imposing a restriction on the form of players' preferences and the version using instead a restriction on the types of blocking coalitions. The use of this strategy in the Japanese case is then described in Section 3, including an analysis of potential alternative national government policies using counterfactual simulations. Section 4 concludes.

2 Theory

Notation follows that of Banerjee, Konishi, and Sönmez [2001]. Specifically, let N be the set of players, and $S \subset N$ a coalition of these players. Π is the set of all possible coalition structures, where a coalition structure $\pi \in \Pi$ is a set of coalitions $\{S_1, \ldots, S_K\}$ such that every player is in exactly one of these coalitions. Suppose that player $i \in N$ has preferences $\preceq_i$ defined over the set $\{S \subset N | i \in S\}$, with $\prec_i$ indicating a strict preference. The extension of these preferences to partitions is easy: if $\pi(i)$ is the coalition that municipality i belongs to in partition π, then $\pi \preceq_i \pi'$ if $\pi(i) \preceq_i \pi'(i)$. Let $\pi \prec_S \pi'$ for some coalition S if $\forall i \in S, \pi \preceq_i \pi'$ and at least one of these preferences is strict. The observed coalition structure is treated as the result of a pure hedonic coalition formation game, where the payoff to each player depends only on the coalition to which it belongs, and not on what other coalitions occur. This is the game introduced by Dreze and Greenberg [1980], except without the possibility of even within-coalition transfers. The inability to negotiate transfers prevents some coalitions from forming:

Example 1. *Let $N = \{1, 2\}$, and u_i be a utility function describing the preferences of player i over coalitions, with*

$$u_1(\{1,2\}) = u_1(\{1\}) + \epsilon_1,$$
$$u_2(\{1,2\}) = u_2(\{2\}) + \epsilon_2. \tag{1}$$

[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.

If $\epsilon_1 > 0$, $\epsilon_2 < 0$, $|\epsilon_1| > |\epsilon_2|$, *then the stable coalition structure is* $\{\{1, 2\}\}$ *if transfers are possible, but* $\{\{1\}, \{2\}\}$ *if they are prohibited.*

Ideally, given a set of preferences, there would exist a unique stable partition: First, the solution set is defined using the von Neumann and Morgenstern [1944] "stable set":

Definition 1. Π^{VNM} *is a stable set with respect to* $(\Pi, <)$ *for some binary operator* $<$ *if*

1. $\nexists \pi, \pi' \in \Pi^{VNM}$ *where* $\pi < \pi'$. *(Internal stability)*

2. $\forall \pi \notin \Pi^{VNM}, \exists \pi' \in \Pi^{VNM}$ *where* $\pi < \pi'$. *(External stability)*

The goal is to define $<$ in a way that is intuitively plausible yet at the same time guarantees that the stable set exists, but this turns out not to be trivial. Consider, for example, the following definition of $<$: $\pi < \pi'$ if $\exists S \in \pi'$ such that $\pi \prec_S \pi'$ and $\forall S' \in (\pi \setminus \pi'), (S' \setminus S) \in \pi'$ or is empty. Unfortunately, with this definition not only is a stable set not guaranteed to exist, but in general it is not possible to devise another plausible method of selecting a single partition as the solution of this type of coalition formation game [Barberà and Gerber, 2007]. The following "roommates problem" illustrates this point:

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

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

With these preferences, no stable partition exists.

Nevertheless, when the Japanese municipalities actually played a coalition formation game, an outcome did occur. The problem is then how to treat observed outcomes such as this one when attempting to estimate parameters. There are at least four ways to proceed: to move to a non-cooperative game structure, to accept set identification rather than point identification, to restrict preferences, or to relax the requirements for stability.

A non-cooperative game is guaranteed to provide a set of equilibrium outcomes, but it is difficult to use in this case as little information is available about the way in which the municipalities actually negotiated, or who made what offers, and so forth. Thus, the specification of the rules of the game would be essentially arbitrary. If the equilibria did not depend on the rules, then the lack of information about the negotiation process would not be important, but it is fairly easy to see that in this sort of coalition formation game, different rules produce different outcomes. For example, if there are a finite number of periods in which a proposer can propose a coalition or coalition structure, then the probability with which various municipalities are selected to be the proposer will change the types of proposals made and accepted. Radically different parameter estimates could be obtained by using different probabilities of having a municipality selected as proposer, and there is no information available on what reasonable proposer weights would be, or even whether the proposer type framework is appropriate.[11]

Another intriguing possibility is that of set identification via moment inequalities, following the framework described by Pakes [2010]. This appears attractive, as the fact that $\{1,2\}$ is observed implies that $\{1\} \prec_1 \{1,2\}$ and $\{2\} \prec_2 \{1,2\}$ under a large variety of equilibrium assumptions. However, in addition to this sort of inequality showing that municipalities do not want to be "too small", in order to identify a bounded set it is also necessary to find inequalities that show that municipalities do not wish to be "too big". This is more challenging than it first appears, as it is true that $\{\{1\},\{2\}\}$ observed implies that either $\{1,2\} \prec_1 \{1\}$ or $\{1,2\} \prec_2 \{2\}$, but it is not clear which of these hold. Designing an error structure that both rationalizes the observed mergers while at the same time allowing estimation using this sort of "either-or" style inequalities is non-trivial, as the most obvious option requires a distributional assumption regarding the error term, thereby eliminating a major advantage of the moment inequalities approach.[12] This, combined with the Ponomareva and Tamer critique regarding model misspecification, are the major reasons why this

[11]These critiques could also be applied to the choice of stability requirements used in the cooperative form game analyzed in this paper. Estimators based on a cooperative form game, however, appear to be computationally simpler to implement than estimators based on a non-cooperative form game. Thus, in the absence of reasons to choose a non-cooperative form game, the paper defaults to the easier cooperative form game.

[12]There is also the possibility of using a model that cannot rationalize all the observed mergers, but it is not clear whether this has significant advantages over a fully-specified model with a sufficiently fat-tailed error term.

paper focuses on a maximum likelihood approach, using either restricted preferences ("RP" from here on) or relaxed stability requirements ("RSR").

2.1 Restricted Preferences

For the RP approach, consider the following restrictions on the form of u_i, the utility that player i derives from a coalition:

$$\begin{aligned} u_i(S) &= u(S) + \alpha_i, \\ u(S) &= v(X_S; \theta) + \epsilon_S, \end{aligned} \tag{3}$$

where v is a function of characteristics X_S of S, taking parameters θ.

The econometrician observes X_S and knows the functional form of v, and the objective is to estimate the parameters θ. The error term ϵ_S is iid of a known distribution. The important restriction here is that if $S \prec_i S'$ then $\forall j, S \prec_j S'$. That is, all agents have identical preferences over coalitions.

Theorem 1 (Farrell and Scotchmer 1988)**.** *If all agents have identical preferences over coalitions, a generically unique stable partition exists.*

Proof. Exactly as given in Farrell and Scotchmer [1988], but repeated in Appendix A because it is a proof by construction, and the algorithm will be used to construct counterfactual partitions later on. □

The restriction on the idiosyncratic error term is strong: it implies that the observed and unobserved characteristics of a coalition are enjoyed equally by all its members. In particular, in the case of municipal mergers it rules out the possibility that a large municipality merging with a smaller neighbour might take advantage of its electoral power within the new amalgamated municipality in order to geographically skew locations of new public facilities. The major benefit of placing this restriction on the error term is that it guarantees uniqueness, and thus estimation does not require any assumption about an equilibrium selection rule.[13]

[13]It might be possible to weaken the restrictions on the error term somewhat by instead assuming the monotonic median voter property [Acemoglu, Egorov, and Sonin, 2008], but the details of this are not immediately obvious.

Suppose that partition π_0 is actually observed. The parameters θ can be estimated via simulated maximum likelihood. The likelihood of π_0 occurring is

$$\begin{aligned}\mathcal{L}(\pi_0 \text{ stable } |\theta) &= \int_\epsilon I(\pi_0 \text{ stable } |\theta, \epsilon) f_\epsilon(\epsilon) d\epsilon \\ &= \int_{\epsilon_0} P(\pi_0 \text{ stable } |\theta, \epsilon_0) f_{\epsilon_0}(\epsilon_0) d\epsilon_0, \end{aligned} \tag{4}$$

where f is the PDF of the idiosyncratic shocks, and ϵ_0 denotes the vector $\{\epsilon_S | S \in \pi_0\}$.[14] This integral can be numerically approximated by taking a set E_0 of random draws of ϵ_0 and calculating

$$\frac{1}{|E_0|} \sum_{\epsilon_0 \in E_0} P(\pi_0 \text{ stable } |\theta, \epsilon_0). \tag{5}$$

Because of the "convenient error partitioning" [Train, 1995] of the above, the probability can be expanded into a product of independent events.[15] If $\mathscr{S}$ is the set of all potential coalitions, as above, then

$$P(\pi_0 \text{ stable } |\theta, \epsilon_0) = \prod_{S' \in \mathscr{S}} P(u(S') < \max_{S \in \text{perp}_{S'}} u(S) | \epsilon_0, \theta), \tag{6}$$

where $\text{perp}_{S'} = \{S | S \in \pi_0, S \cap S' \neq \emptyset\}$ is the set of "perpetrators" necessary to

[14] More formally, define ϵ_1 as the ϵ shocks not in ϵ_0, and note that $f_\epsilon(\epsilon) = f_\epsilon(\epsilon_1, \epsilon_0) = f_{\epsilon_1|\epsilon_0}(\epsilon_1|\epsilon_0) f_{\epsilon_0}(\epsilon_0)$. Then rewrite as follows:

$$\begin{aligned}\mathcal{L}(\pi_0 \text{ stable } |\theta) &= \int_{\epsilon_0} \int_{\epsilon_1} I(\pi_0 \text{ stable } |\theta, \epsilon_1, \epsilon_0) f_\epsilon(\epsilon_1, \epsilon_0) d\epsilon_1 d\epsilon_0 \\ &= \int_{\epsilon_0} \int_{\epsilon_1} I(\pi_0 \text{ stable } |\theta, \epsilon_1, \epsilon_0) f_{\epsilon_1|\epsilon_0}(\epsilon_1|\epsilon_0) f_{\epsilon_0}(\epsilon_0) d\epsilon_1 d\epsilon_0 \\ &= \int_{\epsilon_0} \left[\int_{\epsilon_1} I(\pi_0 \text{ stable } |\theta, \epsilon_1, \epsilon_0) f_{\epsilon_1|\epsilon_0}(\epsilon_1|\epsilon_0) d\epsilon_1 \right] f_{\epsilon_0}(\epsilon_0) d\epsilon_0 \\ &= \int_{\epsilon_0} P(\pi_0 \text{ stable } |\theta, \epsilon_0) f_{\epsilon_0}(\epsilon_0) d\epsilon_0. \end{aligned}$$

This exposition is due to Vadim Marmer.

[15] That is, once the ϵ_0 have been drawn, and thus the $u(S)$ are known for $S \in \pi_0$, the events $u(S') > u(S)$ and $u(S'') > u(S)$ are independent. This conditional independence allows conditional probabilities to be expressed as products of the relevant independent events.

deviate to S'. The likelihood function used for optimization is thus

$$\mathcal{L}(\pi_0 \text{ stable } |\theta) = \frac{1}{|E_0|} \sum_{\epsilon_0 \in E_0} \prod_{S' \in \mathscr{S}} P(u(S') < \max_{S \in perp_{S'}} u(S)|\epsilon_0, \theta). \tag{7}$$

2.2 Relaxed stability requirements

For the RSR approach, suppose that a less restrictive form is imposed on preferences:

$$u_i(S) = v(X_i, X_S; \theta) + \epsilon_{iS} \tag{8}$$

where ϵ_{iS} are iid draws from a known distribution. Here, the utility a player derives from a coalition can depend on interactions between the player's characteristics and those of the coalition, and similarly the ϵ for a given coalition can vary across players. In this case, the existence, but not uniqueness of a stable partition can be guaranteed so long as some restrictions are placed on the types of blocking coalitions that can form. In particular, only two types of potential deviations will be considered when evaluating whether a given partition is stable: refinements, where a subcoalition of a single existing coalition breaks off to form a coalition, and coarsenings, where two or more existing coalitions merge in order to form a new coalition.

To guarantee existence, 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.[16] Because hedonic games are simpler than the "equilibrium coalition structures" that Ray and Vohra examine, refinements and coarsenings will be treated identically. 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

[16] 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.

perpetrators deviated, such that the configuration of perpetrators and residual is stable. More formally, where $\to$ should be read as "blocked by":

Definition 2. *$\pi \to \pi'$ 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 $S = \bigcup Q$ for some $Q \subset \pi$.*

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 $Q' \subset \pi'$, or*
    b) *$\nexists \tilde{Q}$ such that $Q' \to \tilde{Q}$.*

The recursion is well defined since Q' is a proper subset of π'.

Theorem 2. *Let $\twoheadrightarrow$ be the transitive closure of $\to$.*[17] *Then*

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

2. *Π^* is unique.*

3. *Π^* contains a Pareto optimal partition.*

Proof. Straightforward given Ray and Vohra [1997], but provided in Appendix A for completeness. □

All partitions in Π^*, including those that are not Pareto optimal, will be 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.[18]

Estimation of this model is similar to that of the RP model. Following the notation introduced for RP (Section 2.1), let ϵ_0 denotes the vector $\epsilon_{iS}, \forall i \in S, \forall S \in \pi_0$. Suppose that partition π_0 is actually observed. If only partitions in the solution set Π^* are observed, and every partition in Π^* is assumed to be selected with equal probability,

[17]To see why the transitive closure is used here, consider the case where $\pi_1 \nearrow_S \pi_2 \searrow_{S'} \pi_3$. 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 $\to$ because $\pi_1 \not\to \pi_3$.

[18]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, in the Japanese case it is unfortunately computationally infeasible to impose any restrictions that require enumerating the entire stable set.

then the parameters θ can be estimated via maximum likelihood. The likelihood of π_0 occurring is

$$\begin{aligned}\mathcal{L}(\pi_0|\theta) &= \int_\epsilon \frac{I(\pi_0 \text{ stable } |\theta,\epsilon)}{\int_\Pi I(\pi \text{ stable } |\theta,\epsilon) f_\pi d\pi} f_\epsilon(\epsilon) d\epsilon \\ &= P(\pi_0 \in \Pi^*|\theta) \int_\epsilon \frac{1}{\int_\Pi I(\pi \text{ stable } |\theta,\epsilon) f_\pi d\pi} f_\epsilon(\epsilon|\pi_0 \in \Pi^*,\theta) d\epsilon \\ &= P(\pi_0 \in \Pi^*|\theta) E_{\epsilon|\pi_0\in\Pi^*,\theta}\left[\frac{1}{Z}\right], \end{aligned} \tag{9}$$

where Z is the number of stable partitions (*i.e.* $|\Pi^*|$). Since the distribution of ϵ is known by assumption, this could in theory be calculated exactly. Due to computational constraints, however, both terms in the above likelihood function will be estimated. Assume that π_0 in fact consists of $\pi_{0,1}, \pi_{0,2}, ..., \pi_{0,K}$, the outcomes of K independent coalition formation games. Then $Z = Z_1 Z_2 \cdots Z_K$, and take the Kth root of both sides of the above equation. Then

$$\sqrt[K]{\mathcal{L}(\pi_0|\theta)} = \sqrt[K]{P(\pi_0 \in \Pi^*|\theta) E_{\epsilon|\pi_0\in\Pi^*,\theta}\left[\frac{1}{Z}\right]} \tag{10}$$

is a consistent M-estimator for θ. However, for computational reasons it is not possible to calculate the above, so instead a numerical approximation of the expectation will be used. In particular, consider the case where $K \to \infty$. Then

$$\frac{\log Z - \mu_{\pi_0,\theta}}{\sqrt{K}} \sim N(0, \sigma_{\pi_0,\theta}) \tag{11}$$

and

$$\frac{1}{E_{\epsilon|\pi_0\in\Pi^*,\theta}[Z]} \to_p E_{\epsilon|\pi_0\in\Pi^*,\theta}\left[\frac{1}{Z}\right]. \tag{12}$$

Now let

$$E_{\epsilon|\pi_0\in\Pi^*,\theta}[Z] = |\Pi|\rho_{\pi_0,\theta}, \tag{13}$$

where ρ is the fraction of partitions that are stable.[19] ρ will be estimated by randomly selecting a set of partitions Π_A and calculating the probability that they are stable. This results in the estimator

$$\begin{aligned}
&\sqrt[K]{P(\pi_0 \in \Pi^*|\theta)\frac{1}{|\Pi|\frac{1}{|\Pi_A|}\Sigma_{\pi\in\Pi_A}P(\pi_i \in \Pi^*|\pi_0 \in \Pi^*, \theta)}} \\
&= \sqrt[K]{|\Pi|}\sqrt[K]{\frac{P(\pi_0 \in \Pi^*|\theta)|\Pi_A|}{\Sigma_{\pi\in\Pi_A}P(\pi_i \in \Pi^*|\pi_0 \in \Pi^*, \theta)}}. \qquad (14)
\end{aligned}$$

Since $|\Pi|$ does not depend on θ, an equivalent estimator is

$$\sqrt[K]{\frac{P(\pi_0 \in \Pi^*|\theta)|\Pi_A|}{\Sigma_{\pi\in\Pi_A}P(\pi_i \in \Pi^*|\pi_0 \in \Pi^*, \theta)}}. \qquad (15)$$

The actual estimation is performed via numerical approximation of the above probabilities. Specifically, if E_0 is a set of draws from the distribution of ϵ_0, then the approximation is

$$\frac{1}{|E_0|}\sum_{\epsilon_0\in E_0} P(\pi_0 \in \Pi^*|\theta, \epsilon_0)\frac{|\Pi_A|}{\sum_{\pi\in\Pi_A} P(\pi \in \Pi^*|\theta, \epsilon_0, \pi_0 \in \Pi^*)}. \qquad (16)$$

Once again, because of the convenient error partitioning of the above, the probability can be expanded into a product of independent events. Let $\mathcal{S}_0^{\uparrow}$ be the set of all coalitions that could be formed by mergers of the coalitions in π_0, and let $\mathcal{S}_0^{\downarrow}$ be the set of all coalitions that are a subset of a coalition in π_0. Then

$$P(\pi_0 \in \Pi^*|\theta, \epsilon_0) = \prod_{S\in(\mathcal{S}_0^{\uparrow}\cup\mathcal{S}_0^{\downarrow})} P(\pi_0 \not\prec_S S|\theta, \epsilon_0). \qquad (17)$$

[19]Interchanging the expectation and reciprocal operations is valid because

$$\begin{aligned}
E[\frac{1}{Z}] &= E[\frac{1}{e^{\Sigma_{i=K}\log Z_i}}] \\
&= E[e^{-\Sigma_{i=K}\frac{\log Z_i}{K}K}] \\
&= E[e^{-\mu K + o(\frac{K}{\sqrt{K}})}] \\
&= E[e^{-\mu K + o(\sqrt{K})}] \\
&\simeq E[e^{-\mu K}],
\end{aligned}$$

since $\mu > 1$ because $\pi_0 \in \Pi^*$.

Now approximate the denominator by defining ϵ_l and E_l in the same way as ϵ_0 and E_0:

$$P(\pi_l \in \Pi^*|\theta, \epsilon_0, \pi_0 \in \Pi^*) = \frac{1}{|E_l|} \sum_{\epsilon_l \in E_l} \prod_{S \in (\mathcal{S}_l^{\uparrow} \cup \mathcal{S}_l^{\downarrow})} P(\pi_l \not\prec_S S|\theta, \epsilon_0, \epsilon_l, \pi_0 \in \Pi^*). \tag{18}$$

There are two problems with estimating this numerically. First, draws need to be made from $\epsilon_l|\theta, \epsilon_0, \pi_0 \in \Pi^*$, and second, given a draw of ϵ_l from the correct distribution, the required probability needs to be calculated efficiently. Fortunately, for both cases, an application of Bayes' Rule is sufficient. To draw from $\epsilon_l|\theta, \epsilon_0, \pi_0 \in \Pi^*$, first define ϵ_S to be the idiosyncratic shocks to coalition $S \in \pi_l$. If S is not a potential deviation from π_0, then $\epsilon_S|\theta, \epsilon_0, \pi_0 \in \Pi^* = \epsilon_S$ since no additional information is provided by the fact that π_0 is stable. If S is a potential deviation from π_0, then consider the identity

$$\begin{aligned} f(\epsilon_S|\theta, \epsilon_0) = \; & f(\epsilon_S|\theta, \epsilon_0, \pi_0 \not\prec S) P(\pi_0 \not\prec_S S|\theta, \epsilon_0) \\ & + f(\epsilon_S|\theta, \epsilon_0, \pi_0 \prec_S S) P(\pi_0 \prec_S S|\theta, \epsilon_0). \end{aligned} \tag{19}$$

Here $f(\epsilon_S|\theta, \epsilon_0)$ is equal to the unconditional density $f(\epsilon_S)$, which is known by assumption. The second distribution on the right hand side is a set of truncated distributions because if $\pi_0 \prec_S S$ then it must be true that $u_i(S) > u_i(\pi_0)$, and thus

$$\epsilon_i > u_i(\pi_0) - v_i(S), \tag{20}$$

and these can be calculated sequentially. Thus, the desired distribution can be drawn by simulating from

$$f(\epsilon_S|\theta, \epsilon_0, \pi_0 \not\prec_S S) = \frac{f(\epsilon_S|\theta, \epsilon_0) - f(\epsilon_S|\theta, \epsilon_0, \pi_0 \prec_S S) P(\pi_0 \prec_S S|\theta, \epsilon_0)}{P(\pi_0 \not\prec_S S|\theta, \epsilon_0)}, \tag{21}$$

which can be done sequentially for each member of S.[20]

[20] The simplest way of ensuring that π_0 is always in the stable set is to draw a new ϵ for each new proposed $\hat{\theta}$; however, this introduces simulation "chatter", making convergence difficult. Instead of simulating ϵ_0 directly, then, draw quantile indices q_j, and create ϵ_j from q_j fresh for each iteration of $\hat{\theta}$.

The next problem is using these drawn ϵ_l to calculate the probability

$$P(\pi_l \not\prec_S S|\theta, \epsilon_0, \epsilon_l, \pi_0 \in \Pi^*) \tag{22}$$

where S' is some coalition not in π_l. If S' is not a potential deviation from π_0, then the calculation is identical for those done for π_0, described above. However, if S' is a potential deviation from π_0, then the fact that π_0 is stable provides additional information that needs to be taken into account. Consider the following:

$$\begin{aligned} P(\pi_l \not\prec_{S'} S'|\theta, \epsilon_0, \epsilon_l) = \; & P(\pi_l \not\prec_{S'} S'|\theta, \epsilon_0, \epsilon_l, \pi_0 \not\prec S')P(\pi_0 \not\prec_{S'} S'|\theta, \epsilon_0, \epsilon_l) \\ & + P(\pi_l \not\prec_{S'} S'|\theta, \epsilon_0, \epsilon_l, \pi_0 \prec_{S'} S')P(\pi_0 \prec_{S'} S'|\theta, \epsilon_0, \epsilon_1). \end{aligned} \tag{23}$$

Since the left hand side can be calculated and the second term of the right hand side has the same set of truncated distributions described just above with respect to the ϵ_l, rearrangement once again permits calculation:

$$\begin{aligned} & P(\pi_l \not\prec_{S'} S'|\theta, \epsilon_0, \epsilon_l, \pi_0 \not\prec_{S'} S') = \\ & \frac{1 - P(\pi_l \prec_{S'} S'|\theta, \epsilon_0, \epsilon_l) - (1 - P(\pi_l \prec_{S'} S'|\theta, \epsilon_0, \epsilon_l, \pi_0 \prec_{S'} S'))P(\pi_0 \prec S'|\theta, \epsilon_0, \epsilon_1)}{1 - P(\pi_0 \prec_{S'} S'|\theta, \epsilon_0, \epsilon_l)}. \end{aligned} \tag{24}$$

Everything on the right hand side of this equation can be computed quickly, making optimization feasible.

3 Application

Treating municipal mergers as a pure hedonic coalition formation problem is consistent with anecdotal evidence concerning how mergers are effected. Negotiations regarding compensation seem to be rare, even though controversy is common and the results of unrest sometimes significant.[21] Some of the involved municipalities may be in favour of a proposed merger while others may be opposed, but those in favour do not seem to promise large transfers to those opposed in order to secure their cooperation. This suggests that there is some problem with contractibility in polit-

[21] For example, in Canada the merger of all municipalities on Montreal Island in 2002 was a major cause of the provincial government losing the next election, and the demerger process following that election meant that on net only limited change occurred despite significant economic and political costs.

ical mergers such that transfers are difficult or impossible, and thus it seems more plausible to model mergers as a coalition formation game without transfers. First, a model of jurisdictions with heterogeneity in individual ideal points is presented, in the style of Greenberg and Weber [1986], Demange [1994], and in particular Alesina and Spolaore [1997]. The Japanese data used is then described, and parameters are estimated via the methods presented in the preceding section.

3.1 Municipal Public Goods Model

Suppose that each municipality m provides a public good of quality q_m at a cost of $q_m \cdot c(P_m)$, where P is population, and c exhibits economies of scale. The municipality levies taxes at rate τ_m and receives transfers T_m from the national government. Population is distributed on a plane, and the public good is provided at a single physical location x_m on this plane. Suppose that there is some minimum tolerable level of public good provision $\underline{q}$. Individual utility will be assumed to be additively separable:

$$u_i(q_m, \tau_m, \theta_m) = \beta_0 \log((1-\tau_m)y_i) + \beta_1 \log(q_m - \underline{q}) + \beta_2 \ell_i(\theta_m) + \epsilon_{im} \tag{25}$$

where ℓ_i is the distance between individual i and location of the public good, with $\beta_2 < 0$. The error term ϵ is irrelevant until the possibility of mergers is considered. The location of the public good is a multidimensional political decision, a problem which has no generally accepted solution concept. However, if this decision is made by a single elected official, such as a mayor, and voting in elections is determined via a probabilistic voting model where vote probabilities are linear in utility difference between two candidates, then the policy chosen will in fact be the socially optimal policy [Banks and Duggan, 2005].[22] In this case, this will set θ_m to the location of the generalized median voter.

The optimal quality supplied is determined by the budget constraint $\tau_m Y_m = q_m c(P_m) - T_m$, where Y_m is total income of all individuals in the municipality, which

[22] 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.

leads to

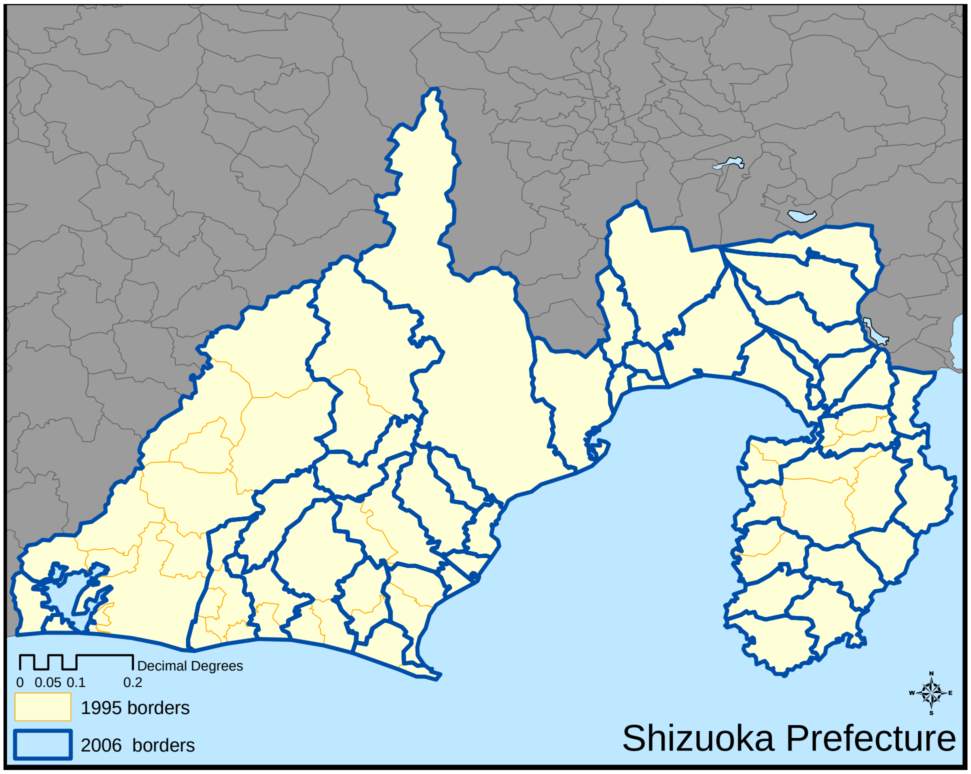
$$q_m^* = \beta_1 \frac{Y_m + T_m - \underline{q}c(P_m)}{c(P_m)} + \underline{q} \tag{26}$$

$$\tau_m^* = 1 - \beta_0 \frac{Y_m + T_m - \underline{q}c(P_m)}{Y_m}, \tag{27}$$

if $\beta_0 + \beta_1 = 1$. Due to the functional form assumption all individuals prefer τ_m^* regardless of income, and thus this is the tax rate implemented.

In this model, individuals do not move or otherwise change their ideal point and population is assumed to be constant. In reality, residence choice is endogenous to government characteristics, as discussed in the literature established by Tiebout [1956] and others. This endogeneity is not included in this paper for three reasons. First, mobility is lower in Japan than in most other developed countries and a large portion of the inter-municipality moves reported in the census appear to be temporary. Endogenous relocation is thus less of a concern than in other countries. Second, there is no evidence of tax competition. The majority of municipalities charge a standard tax rate and even though municipalities are allowed to set a different rate (within a band), few choose to exercise this option. This is consistent with the model presented above, and combined with the national government transfer scheme results in the endogeneity problem being less severe than it would be in other contexts. Finally, from an implementation perspective, endogeneity would result in future population and other characteristics of a municipal merger depending on what other mergers occurred in the surrounding area. This would change the nature of the coalition formation game from a characteristic function game to a partition function game, which is substantially more computationally intensive to estimate, and likely infeasible without further theoretical or technological developments. Population dynamics are thus ignored, with the estimates that follow focusing on a single period game with player characteristics determined based on current government data sources.[23]

Figure 1: Shizuoka Prefecture



3.2 Data

There were 3,255 municipalities in 1999 at the start of the merger period, divided into 47 prefectures (similar to U.S. states). Since mergers do not cross prefectural boundaries, each prefecture is treated as a separate coalition formation game.[24] Figure 1 shows the mergers that occurred in Shizuoka Prefecture. Mergers were voluntary, and needed to be approved by the municipal council of every participating municipality.[25]

[23]Similar results were obtained when 2005 census data is used instead of the 1995 data currently used. Predicted future municipal population could be created based on census data, but the similarity of results obtained with data from different census years suggests that this exercise may not produce particularly interesting results.

[24]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.

[25]In about a third of cases, referenda were held. Nominally, these were 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.

The involvement of the municipal council suggests that some form of legislative bargaining model may be needed; however this paper abstracts away from this issue and supposes that a single elected official such as the mayor is responsible for decision-making. In the simplified framework presented here, these decisions will be optimal from the perspective of municipal residents, and parameters to the utility function given in Equation 25 can thus be estimated by examining the mergers that actually occurred.[26]

The number of municipalities dropped to 1,750 in 2010, due to 598 mergers ranging in size from two municipalities up to fifteen municipalities. These mergers occurred mainly due to changes to the rules determining transfers T. Historically, transfers were determined by the formula

$$T_m = \max(\tilde{c}_m(P_m) - .75\bar{\tau}Y_m, 0), \tag{28}$$

where $\tilde{c}_m$ is the estimated cost to the municipality of providing those services (referred to as "Standard Fiscal Need" in official documents), and $\bar{\tau}$ a reference tax rate set by the national government. Here the scale for quality has been normalized such that the "national minimum" level of quality used by the government to compute Standard Fiscal Need is 1. Although there were provisions for municipalities to merge, there was little incentive for them to do so, because if a coalition S decided to form a new (amalgamated) municipality, T_S would be calculated exactly as in Equation 28:

$$T_S = \max(\tilde{c}_S(P_S) - .75\bar{\tau}Y_S, 0). \tag{29}$$

Thus almost all savings would be passed to the national government, and even a slight preference for smaller jurisdictions ensured that residents would be opposed to mergers.[27]

During the fiscal difficulties of the early 1990s, the 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. These reforms consisted mainly of a lump sum cut in transfers to all municipalities, making

[26] The possibility of future mergers is ruled out. Given that the previous set of municipal mergers occurred in the 1960s, experience suggests that any future mergers are likely far enough away that a reasonable discount rate reduces their importance to the point where they can safely be ignored.

[27] 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.

it difficult for small municipalities to maintain service quality unless they merged, and change to Equation 29 to reduce the penalty associated with merging.

Data sources for municipal population, income per capita, etc. are described in Appendix B. Details regarding $\tilde{c}$ and the changes to the transfer formula are given in Appendix C. Further general information on the Japanese local public finance system available from Mochida [2008]. Hayashi, Nishikawa, and Weese [2010] examine the issues surrounding c in some detail and conclude that $\tilde{c} \simeq c$: that is, the national government estimates accurately reflect the actual efficiencies of scale in the provision of public services.[28] These efficiencies of scale are substantial, as shown in Figure 2 (see Appendix C for details).

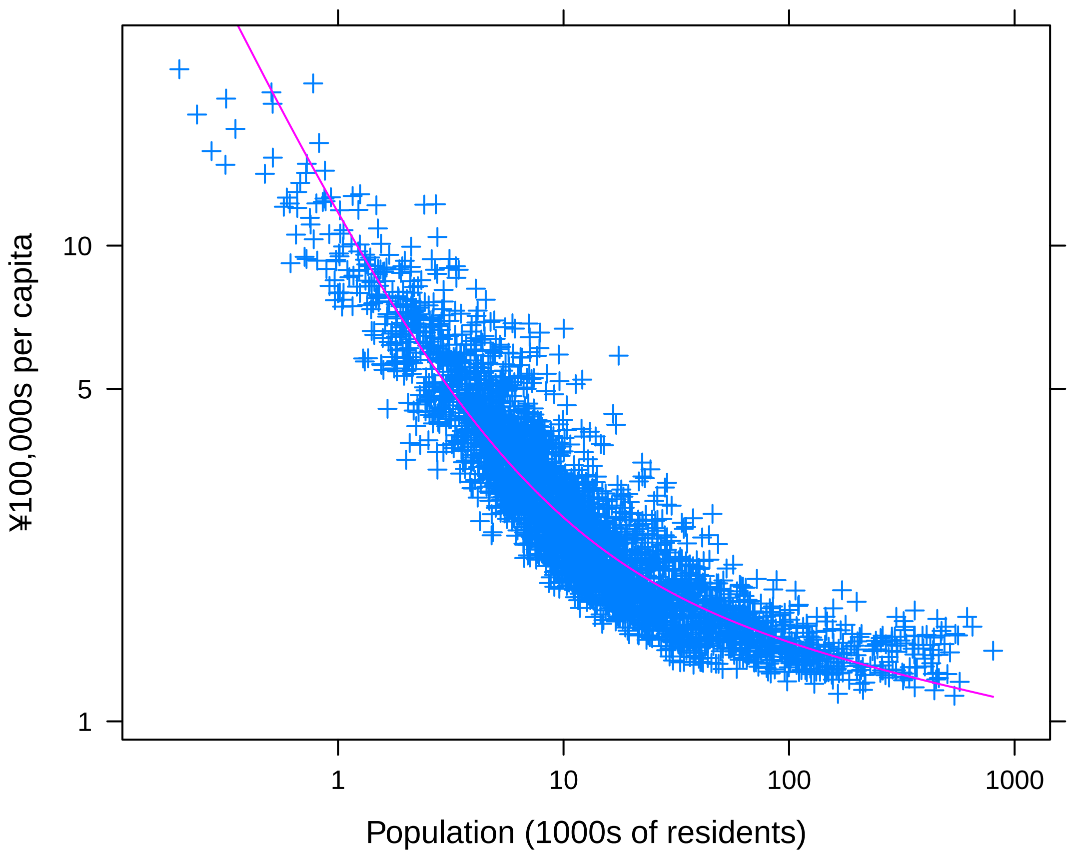
A final interesting feature of municipal finance in Japan is that there is little variation in tax rates charged by municipalities: despite the fact that local governments have the authority to choose any rate within a band established by the national government, almost all charge $\tau_m = \bar{\tau}$. One possible explanation is political: municipalities charge $\bar{\tau}$ even though $\bar{\tau} > \tau_m^*$ for most m because of pressure from higher levels of government or the risk of negative publicity. This is unsatisfying in that it suggests that a much more complicated model, one involving political processes about which little is known, is necessary in order to explain the observed municipal tax system. Another possibility is to set $\underline{q} = 1$, that is, to make the official national minimum standard the boundary between intolerably low and acceptable levels of public services for individuals. This is also unsatisfying in that it implies that any level of government service below the national minimum would result in infinite disutility; however, it has the major advantage that Equation 27 reduces to

$$\tau_m^* = 1 - \beta_0(1 - .75\bar{\tau})$$

for municipalities that are receiving transfers, using Equation 28 and the assumption that $\tilde{c}_m = c_m$. Therefore, $\tau_m^* = \bar{\tau}$ if $\beta_0 = \frac{1-\bar{\tau}}{1-.75\bar{\tau}}$, which is about 0.97 when $\bar{\tau} = 0.12$,

[28]To briefly summarize Hayashi et al. [2010], over-representation of rural areas within the Japanese Diet means that the risk is that government estimates overstate efficiencies of scale, rather than understating them. If the true efficiencies of scale were lower than the government estimates shown in Figure 2, then smaller municipalities were receiving rents. If such rents existed, then either they were distributed to residents, which should be visible in land prices, or they were captured by a group of political insiders, which should result in rural political leaders being more opposed to municipal mergers than their constituents; however, neither of these phenomena are observed.

Figure 2: Standard Fiscal Need



the observed average tax rate. Although this value for β_0 is higher than the estimated value discussed later ($\hat{\beta}_0 = 0.91$), bootstrap analysis suggests that the null hypothesis that $\beta_0 = \frac{1-\bar{\tau}}{1-.75\bar{\tau}}$ should not be rejected. The assumption that $\underline{q} = 1$ will thus be used both for estimation and the analysis of counterfactual policies.

3.3 Estimation

Consider Equation 25, but simplify and replace municipality m with coalition S:

$$u_i(q_S, \tau_S, \theta_S) = \beta_0 \log(1 - \tau_S) + \beta_1 \log(q_S - \underline{q}) + \beta_2 \ell_i(\theta_S) + \epsilon_{iS}. \tag{30}$$

The local politician is assumed to choose the social optimum, as discussed above. As there is no private information, the choice of q^*_S, τ^*_S, and θ^*_S can be predicted for any coalition S. Thus, with some abuse of notation, the preferences for the politician from municipality m can be described by

$$u_m(S) = \beta_0 \log(1 - \tau^*_S) + \beta_1 \log(q^*_S - \underline{q}) + \beta_2 \frac{\sum_{i \in m} \ell_i(\theta^*_S)}{P_m} + \epsilon_{mS}, \tag{31}$$

where $\epsilon_{mS} = \frac{1}{P_m} \sum_{i \in m} \epsilon_{iS}$ and might plausibly be normally distributed due to the Central Limit Theorem. For technical reasons, however, the assumption will be that ϵ_{mS} is drawn from an Extreme Value Type II distribution.[29]

For estimation via the RP method, the restriction $\epsilon_{mS} = \epsilon_{m'S} = \epsilon_S$ is also required. Furthermore, the distance term $\frac{1}{P_m} \sum_{i \in m} \ell_i(\theta^*_S)$ cannot be used as is with this method, as it differs between members of the same coalition S. Thus, for the RP approach, $\bar{\ell}$, the mean distance over all individuals in the coalition S, will be used in place of the actual ℓ_i values for individuals within each municipality. This is equivalent to assuming that all individuals within S will be randomly reassigned a different individual's ideal point after S forms, which is not plausible. This necessary but undesirable assumption will bias both the estimate of β_2 and other parameters, as discussed in the following section. The RSR method allows arbitrary utility functions, and thus the distance term can be used as given in Equation 31.

The determination of $\mathscr{S}$, the set of potential alternative mergers that need to be checked during estimation, is slightly more problematic. There are a number of large mergers observed, with the largest involving fifteen municipalities. Almost all observed mergers are geographically contiguous. However, even after restricting $\mathscr{S}$ to contiguous coalitions of size fifteen or less, there are still over 10^{16} possibilities, which is computationally infeasible.[30] Most of these coalitions, however, look very

[29]Specifically, if ϵ is drawn from an EV Type II ("Frechet") distribution, then the log probability that $S \prec_m S'$ given ϵ_{mS} simplifies to $-(v_{mS'} - u_{mS})^{-\alpha}$, where α is a shape parameter and v is the non-idiosyncratic part of u. This both substantially speeds calculation and helps to avoid floating-point underflow errors, which are potentially problematic when performing non-trivial calculations on large numbers of probabilities close to both 0 and 1. The Frechet distribution may appear undesirable because it is bounded on one side; however, preferences of m over S and S' depend on the difference between ϵ_{mS} and $\epsilon_{mS'}$, which is unbounded. Any observed outcome can thus be rationalized under this distributional assumption.

[30]More specifically, there are thirteen observed mergers that are not geographically contiguous, usually because one of the participants dropped out late in the merger process. Although the law stated that mergers were to be contiguous, these exceptions were allowed. Islands with only a single

different than the actually observed coalitions. In particular, they tend to be a thin line of municipalities, stretching almost all the way across a prefecture. On average, individuals in these coalitions would have very high distance ℓ, and the coalitions are thus not likely to form. Including this group of "low probability" coalitions in the calculations would be ideal. However, for computational reasons these will instead be ruled out through the use of a restriction: coalitions will not be allowed to cross more than two county boundaries, using county definitions from the Meiji era.[31] This restriction dramatically reduces the number of large coalitions that need to be considered: with fifteen-municipality coalitions, only one coalition in a billion involves three or fewer counties.[32] This reduces the total number of alternatives that need to be considered to about 20 million per prefecture, which is computationally feasible.

Another problem that is only relevant to the estimation via RSR is the estimation of the size of the stable set. Since the number of partitions grows exponentially with the number of municipalities, it is not possible to examine all partitions. The total number of partitions is unknown but bounded above by the Bell numbers (Sloane #A000110), which are greater than 10^{100} for larger prefectures such as Hokkaidō. Therefore, a random sample is drawn instead. It is, however, non-trivial to randomly sample from a set which is too large to be enumerated. Thus, random draws are obtained using Markov chains, as described in Appendix D.

3.4 Results

The results are shown in Table 1, with scale determined by the restriction $\beta_0 + \beta_1 = 1$ and standard errors approximated via likelihood ratio tests. Columns I and II are

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. There are, however, two cases in which municipalities on an island merged with municipalities on the mainland other than the closest one. There are also six cases where municipalities on two separate islands merged together. Thus, about 3.5% of mergers (21/588) are not contiguous. No additional mergers violating contiguity are generated as comparison coalitions, although the mergers did occur are retained in the observed partitions, and may also appear in alternative partitions.

[31] 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.

[32] 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.

Table 1: Dependent variable is $u_m(S)$, utility to municipality m from merger S

	relaxed stability requirements		restricted preferences	
	I	II	III	IV
CONSUMPTION (β_0)	0.909	0.913	1.014	1.020
	(0.074)	(0.098)	(0.120)	(0.127)
GOVERNMENT (β_1)	0.091	0.087	−0.014	−0.020
	(0.002)	(0.005)	(0.001)	(0.001)
DISTANCE ($10 \cdot \beta_2$)	−0.086	−0.072	−0.035	−0.039
	(0.001)	(0.003)	(0.001)	(0.001)
DISTANCE_POL		−0.068		0.015
		(0.024)		(0.006)
scale	1.955	0.210	0.152	0.275
shape	107.569	16.281	9.150	14.880
N (prefectures)	46	46	40	40

estimates using the RSR approach, while columns III and IV are estimates using the RP approach. Columns I and III use exactly the utility function discussed previously, while Columns II and IV include one additional variable: DISTANCE_POL is the average political distance to the new median voter on a one dimensional political spectrum, using political spectrum location estimates from Kubo, Mitsuyo, and Kishimoto [2006]. Results obtained via the RSR approach are consistent with theory: higher private (post-tax) consumption and better government services are desirable, while greater distance, both geographic and on the political spectrum, is undesirable.

Assuming that income and transfers per capita are identical across all municipalities, and ignoring the effect of the transfer scheme on merger incentives, the ratio of the coefficients regarding geographic distance and government services in Column I of Table 1 imply that a municipality would be willing to accept an increase in average distance between residents' ideal points and policy location of 10km (disutility of 0.086), in exchange for decreasing the per capita cost of providing government services by a factor of e (utility of 0.091). From the perspective of private consumption, an individual with average income is willing to accept about ¥25,000 per year in exchange for an increase in distance of 1km. This estimate may seem high; however, municipalities are responsible for public school facilities, and municipal mergers are often followed by school amalgamations. For many families, travel to the public facilities in question occurs more days than not.

If a social planner were drawing initial borders on a featureless plane with uniform population density equal to the average population density of Japan (about 340 per sq. km), the optimal population size for a set of uniformly sized municipalities would be the value of P_m that maximizes

$$\beta_0 \log(\frac{Y_m - c(P_m)}{Y_m}) + \beta_1 \log(\frac{Y_m - c(P_m)}{c(P_m)}) + 0.377\beta_2\sqrt{P_m/340}, \qquad (32)$$

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. This gives an optimal size of about 80,000 residents, which is below Japanese estimates of the efficient size for a municipality: The Ministry of Internal Affairs' "Standard Municipality" has a size of 100,000, and the Ministry's target of 1,000 municipalities implies a target population (on average) of 125,000.[33] This difference may be due to the fact that the model used in this paper does not include the ability of "designated cities" to assume some responsibilities that would normally be handled by the prefecture. This appears to have been one of the incentives for very large mergers, such as Niigata City (size fifteen) or Hamamatsu City (size eleven), and the model underpredicts the number of these very large mergers.[34]

While the estimates obtained using the RSR approach match theoretical predictions, those obtained by the RP approach (Columns III and IV) do not. In particular, the coefficient on public services is negative, which suggests that the model is misspecified. A key difference between the RSR approach and the RP approach, as discussed in the previous section, is that the DISTANCE variable for municipality m in coalition S is based on the geographic distance $\ell_i(\theta_S^*)$ for individuals i in municipality m in the former case, but all individuals in the entire coalition S in the latter case. The problem with the averaging in this latter case is that it understates the disadvantage of merging for relatively smaller municipalities, and overstates it for relatively larger municipalities. In the extreme, if each municipality had its population at a single point, there would be no change in distance for the larger of a pair of

[33]Hayashi [2002] finds that the smallest city of efficient scale has a population of 120,000, based on third-party ratings of municipal service quality.

[34]Adding the possibility of moving certain public goods from the prefectural level to the municipal level would substantially complicate the model, as it would imply that there are multiple public goods displaying different efficiencies of scale. This, combined with the fact that there are relatively few "designated cities", is why this feature is not included in the model.

Table 2: Monte Carlo simulations (median estimate and 95% interval)

	true value	relaxed stability requirements		restricted preferences
			(rest. dist.)	
	I	II	III	IV
CONSUMPTION (β_0)	0.913	0.908	0.970	1.010
		[0.245, 0.981]	[0.820, 0.995]	[1.003, 1.019]
GOVERNMENT (β_1)	0.087	0.092	0.030	-0.010
		[0.019, 0.755]	[0.005, 0.180]	[-0.019, -0.003]
DISTANCE ($10 \cdot \beta_2$)	-0.072	-0.067	-0.027	-0.015
		[-0.497, -0.017]	[-0.119, -0.011]	[-0.033, -0.010]
DISTANCE_POL	-0.068	-0.067	-0.012	-0.010
		[-0.448, -0.008]	[-0.088, 0.013]	[-0.038, 0.020]
scale	0.210	0.204	0.055	0.093
shape	16.281			
N (pref. per rep.)	15			
R (replications)	100			

municipalities if they merged, as the location of the generalized median voter would remain unchanged. An inaccurately low estimate of the value of government services counteracts this effect, as the change in the quality of government services is larger for the smaller municipalities participating in a merger than for the larger ones.

Table 2 shows this effect with artificial data created from the parameter values shown in Column I, which correspond to those in Column II of Table 1, and are based on the RSR model. Column II of Table 2 verifies that these parameters can be recovered with acceptable finite sample bias and simulation noise via the strategy presented in this paper. Column IV of Table 2 shows that if this simulated data is used to estimate parameters based on the RP model, a negative coefficient on government services is obtained, as in Columns III and IV of Table 1. Column III of Table 2 shows results obtained by running the RSR model using the modified distance data necessary to meet the restrictions of the RP model. The parameter estimates in Column III are closer to those of Column IV than Column II, showing that the majority of the bias in the estimates in Column IV is due to the different definition of distance that must be used, rather than the different distributional assumption inherent in the model.[35] For the counterfactual policy simulations presented below, then, the RSR model will be treated as the correct model.

A major advantage of having coefficient estimates for a structural model is the ability to conduct counterfactual analysis. Two alternative national government policies can be examined: first, national government enforcement of transfers negotiated between municipalities during the coalition formation process; second, an incentive scheme for richer municipalities that participate in mergers.

3.5 Counterfactual simulations: Transferable utility

Suppose that the central government offered to enforce whatever transfers resulted from decentralized negotiations amongst municipalities. That is, if municipality m promised to make a certain transfer from its current residents to the current residents of municipality m', the central government would ensure that this was actually carried out. Assume that the transfers negotiated are "small", in the sense that a linear approximation of utility around transfers of zero is reasonable:

$$\begin{aligned} u_m(S,z) &= \beta_0 \log((1-\tau_S)y_m + \frac{\sum_{m'\in S} z_{mm'}}{P_m}) + ... + \epsilon_{mS} \\ &\simeq \beta_0 \log((1-\tau_S)y_m) + \frac{\beta_0 \sum_{m'\in S} z_{mm'}}{(1-\tau_S)y_m P_m} + ... + \epsilon_{mS}. \end{aligned} \tag{33}$$

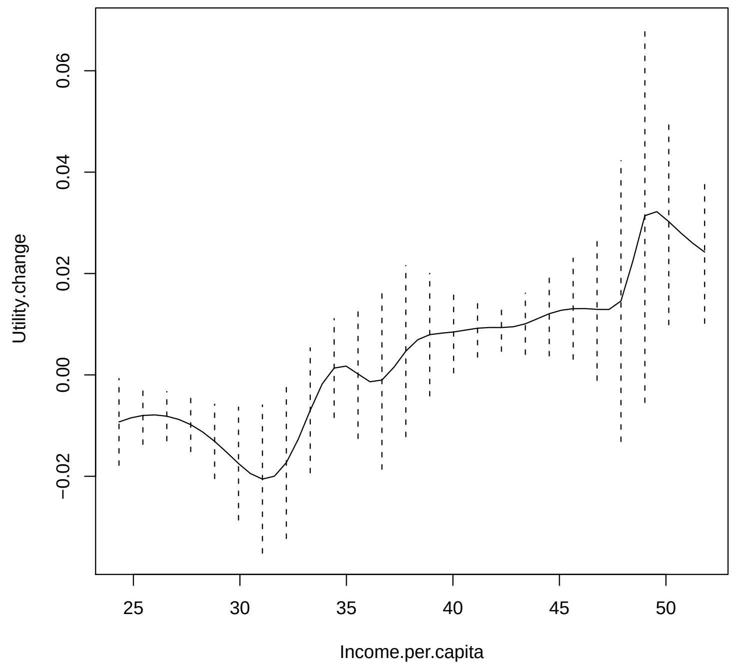
Where $\sum z$ is total transfers received from other municipalities. Now define

$$\begin{aligned} \tilde{u}_m(S,z) &= \frac{(1-\tau_S)y_m P_m}{\beta_0} \cdot u_m(S) \\ &\simeq \sum_{m'\in S} z_{mm'} + \beta_0 log((1-\tau_S)y_m) + ... + \tilde{\epsilon}_{mS}, \end{aligned} \tag{34}$$

where $\tilde{\epsilon}_{mS} = \frac{(1-\tau_S)y_m P_m}{\beta_0}\epsilon_{mS}$. Here, $\tilde{u}$ approximates a standard transferable utility cooperative game. This does not have a unique solution, but is covered by the Ray and Vohra [1997] approach detailed above, and so will result in some sort of stable

[35] The fact that distributional assumptions, which are arbitrary, do play a role in the estimates is of course a concern. Currently, the idiosyncratic component of preferences for municipality m regarding coalition S are assumed to be either uncorrelated with the idiosyncratic component of preferences for municipality m' regarding coalition S (in the RSR model) or completely correlated (in the RP model). Allowing more flexible correlation structures would be an obvious improvement, but computational requirements would likely increase substantially.

Figure 3:



set. If random ϵ are drawn, then some idea of the efficiency gain of enforcing the TU game, versus simply administering a fixed incentive structure, can be obtained. The exact transfers, however, depend on which stable coalition structure forms and exactly how the surplus from each coalition is divided. For a given coalition S, the possible utilities of the municipalities are determined by the possible values of x, where the following conditions are satisfied:

1. $\sum_{m \in S} x_m = V(S)$, and
2. $\sum_{m \in S'} x_m \geq V(S') \quad \forall S' \subset S$,

where $V(S) = \sum_{m \in S} \tilde{u}(S)$ is the value of coalition S.

It seems extremely unlikely that poor municipalities would have more bargaining power than rich municipalities, and thus an "equitable" distribution of the surplus seems to be the most optimistic scenario. A more pessimistic scenario would give most of the bargaining power to richer municipalities, with a resulting increase in post-merger inequality.

The nucleolus is used as a "best case" equitable division: it is the allocation which maximizes over all potential deviating coalitions S' the smallest difference between the amount allocated to the members of S' and that which they could obtain if they deviated.[36] To estimate the utility obtained in this case, 100 separate draws of ϵ were performed for the eight smallest prefectures, and the resulting nucleoli were averaged. This was compared to the no-transfer case, using the same 100 draws of ϵ. The results are shown graphically in Figure 3. Municipalities with a per capita income of less than ¥3,500,000 (about 18% of total population) are worse off, relative to the no-transfers case, while richer municipalities are much better off. Thus, making transfers feasible may or may not be optimal for the national government, depending on its social welfare function. A worst-case scenario was also examined, where richer municipalities are assumed to have all of the bargaining power, and thus as much surplus as possible is transferred to the richest municipality, and then the next richest. The results are very similar to those shown in Figure 3, although the regressive nature of the transfers is slightly more pronounced.

3.6 Counterfactual simulations: Incentives to merge

If transfers between players are disallowed or otherwise impossible, but there is a social planner that can provide incentives for coalition formation, then at least in the two player case, the social planner should offer such incentives:

Example 3. *In the setup described in Example 1, if ϵ_1 and ϵ_2 are random variables with density f, then before the ϵ are known, $\exists \tau, \delta > 0$ such that both players would be in favour of a social planner offering an incentive δ for coalition formation:*

$$\tilde{u}_i(\{1\}) = u_i(\{1\}) - \tau$$
$$\tilde{u}_i(\{1,2\}) = \tilde{u}_i(\{1\}) + \epsilon_i + \delta$$
$$E[1(\tilde{u}_1(\{1,2\}) > \tilde{u}_1(\{1\})) \cdot 1(\tilde{u}_2(\{1,2\}) > \tilde{u}_2(\{2\})) \cdot \delta] = \tau \tag{35}$$

[36] More specifically, x is determined by $\underset{x}{\operatorname{argmin}} \max_{S' \subset S} E(S')$, where $E(S')$ is the excess of coalition S': $E(S') = V(S') - \sum_{m \in S'} x_m$. Note that if S is part of a stable partition, $E(S') < 0 \quad \forall S' \subset S$. There may be more than one x that leads to the same maximum above, but by continuing to minimize excesses lexicographically, a unique value of x is obtained.

Ex ante, both players are in favour of this incentive being offered because

$$E[u_i] = E[u_i(\{i\})] + E[1(u_1(\{1,2\}) > u_1(\{1\})) \cdot 1(u_2(\{1,2\}) > u_2(\{2\})) \cdot (u_i(\{1,2\}) - u_i(\{i\}))] \quad (36)$$

and thus

$$\frac{\partial E[u_i]}{\partial \delta}|_{\delta=0} = f(0)(E[\max(u_i(\{1,2\}), u_i(\{i\}))] - u_i(\{i\})) > 0 \quad (37)$$

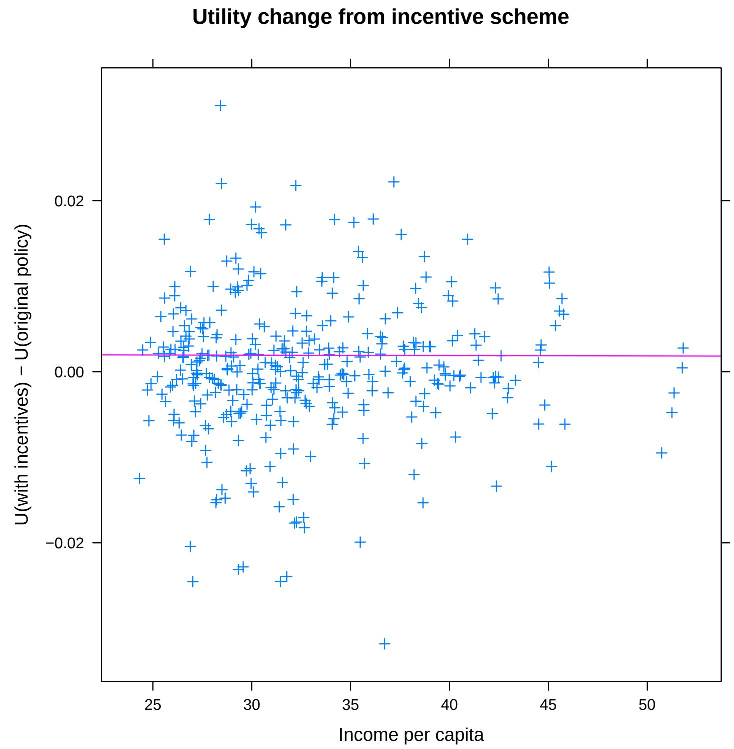
Intuitively, the gains from the incentive are first order, occurring whenever one player is almost indifferent between forming the {1,2} coalition or not, and the other player is very much in favour. The losses, on the other hand, are second order: when both players are almost indifferent, the incentive results in a {1,2} coalition now forming when it wouldn't have previously. In a game with more than two players, however, any new coalition that forms due to this sort of incentive might displace an existing non-singleton coalition, thus changing the set of stable coalition structures. This is an additional cost of offering the incentive, and any general theorem regarding the efficiency of offering incentives to form (non-singleton) coalitions would have to have restrictions to ensure that the cost of such a disruption of existing coalitions is not too high. The definition of "too high" in this case would depend on the social welfare function used, but a simulation shows that welfare improving policies do exist for reasonable social welfare functions.

In particular, suppose that the national government offered an additional, budget balanced incentive for certain municipalities to merge. The targeted municipalities should be those that are most likely to be opposed to mergers that would benefit other municipalities, and the most likely municipalities to fall into that category are richer municipalities. Consider the policy that offered a transfer equivalent to 0.3% of income, to residents of richer municipalities that participated in a merger, where a "richer" municipality is defined as one that had income per capita above the average for the merger that they are participating in. This transfer would be paid for by an increase in the income tax on everyone.

To determine the effects of such a policy, stable partitions are simulated under both the actual and the proposed alternative policies.[37] The difference between the

[37]These partitions are produced by generating random coarsenings of an all-singleton partition

Figure 4:



actual and alternative policies are shown in Figure 4.[38] In addition to increasing utility overall, the increase is the same for poorer and richer municipalities. This is in contrast to the case with negotiated transfers, where poorer municipalities saw decreased utility. For some set of inequality-averse social welfare functions, a social planner might choose not to allow municipalities to negotiate transfers, but might instead set a fixed incentive scheme, one that rewarded desirable municipalities for merging with their neighbours. Such an incentive scheme, however, suffers from the problem that it appears to be regressive, in that it taxes everyone in order to make transfers to the rich, and thus might be infeasible for political reasons.

and then drawing from this set of partitions after all the duplicates have been removed. This is a consistent method of drawing a sample of stable partitions with uniform selection probability, although it is biased (i.e. the probability of selection is not uniform) when the random coarsenings drawn do not enumerate all the stable partitions. The degree of bias depends on the number of random coarsenings drawn, but this bias does not appear to be important, since changing the number of random coarsenings generated does not change the results.

[38]The regression line shown is weighted by municipal population.

4 Conclusion

This paper estimated the parameters determining preferences in a cooperative form political coalition formation game, using two different sets of assumptions and definitions of the solution set. The results obtained via the RSR approach are consistent with intuition, while some coefficient estimates obtained via the RP approach are implausible. This difference is due mainly to the fact that preferences over coalitions differ by player, and this can be incorporated within the RSR approach. The parameters estimated via the RSR approach are used to examine potential alternative national government policies. Counterfactual simulations suggest that an alternative incentive scheme that rewarded relatively rich municipalities for merging would have resulted in welfare improvements under most reasonable social welfare functions. Allowing transfers to be negotiated between municipalities may or may not be superior, depending on the national government's aversion to inequality and the bargaining power of the various municipalities. The latter is likely unknown to the national government, and thus even if transfers between municipalities could be enforced, it may not be beneficial for the national government to do so.[39] At least as important as the implications to government policy, however, is the methodology developed. A coalition formation game without transfers accurately describes many real-world phenomena, but it is rarely estimated in the empirical literature. As the price of computing power decreases, however, the number of uses of this sort of model that are feasible should increase. Although the game presented in this paper could be estimated only because the geographical nature of the data permitted a large number of possible coalitions to be discarded, such restrictions are less likely to be necessary in the future. The results given above, then, are hopefully only an early indication of the applications of this type of coalition formation model.

[39] One possibility that was not considered in this paper is that of a tax on negotiated transfers. In the simplest price control model, a tax redistributed to the consumer should be able to mimic a price control, but with the assurance that the consumers with the highest willingness to pay obtain the good. To the extent that the inability to make transfers is like a price control at zero, then, it could be that the optimal policy for the government – rather than specifying a fixed incentive scheme to encourage rich municipalities to merge with their neighbours – would be to allow transfers, but tax them heavily and redistribute the revenue obtained to the poorest municipalities. Overall, the problem bears some resemblance to the classic rent control problem.

References

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

Daron Acemoglu, Georgy Egorov, and Konstantin Sonin. Dynamics and stability of constitutions, coalitions, and clubs. Technical report, NBER, August 2008.

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

Mark Armstrong and John Vickers. A model of delegated project choice. *Econometrica*, 78(1):213–244, 2010.

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.

David M. Brasington. Joint provision of public goods: the consolidation of school districts. *Journal of Public Economics*, 73(3):373–393, 1999.

B. Dahlby and L.S. Wilson. Fiscal capacity, tax effort, and optimal equalization grants. *The Canadian Journal of Economics*, 27(3):657–672, 1994.

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

Klaus Desmet, Michel Le Breton, Ignacio Ortuno-Ortin, and Shlomo Weber. The stability and breakup of nations: A quantitative analysis. *SSRN eLibrary*, June 2009.

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.

Joseph Farrell and Suzanne Scotchmer. Partnerships. *The Quarterly Journal of Economics*, 103(2):279–297, May 1988.

Frank Flatters, Vernon Henderson, and Peter Mieszkowski. Public goods, efficiency, and regional fiscal equalization. *Journal of Public Economics*, 3(2):99–112, May 1974.

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. Congestion, technical returns and the minimum efficient scales of local government expenditures: The case of japanese municipalities. Technical Report 01-01, Institute for Research in Bus. and Econ., Meiji Gakuin, 2002.

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"].

Hideki Kido and Etsuhiro Nakamura. Strategy and environment: Determinants of municipal amalgamations. *Annals of the Japanese Society for Public Administration*, 43:112–130, 2008. [In Japanese].

Takahiro Kubo, Yamamoto Mitsuyo, and Kazuo Kishimoto. Estimating the downsian locations of parties from the numbers of votes gained by the parties. *Transactions of the Japan Society for Industrial and Applied Mathematics*, 16(4):563–574, December 2006. [In Japanese].

Antonio Merlo. Whither political economy? theories, facts and issues. In Richard Blundell, Whitney K Newey, and Torsten Persson, editors, *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, number 41 in Econometric Society monographs. Cambridge University Press, Cambridge, 2006.

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.

Nobuki Mochida. *Fiscal Decentralization and Local Public Finance in Japan*. Number 107 in Routledge frontiers of political economy. Routledge, London, 2008.

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.

Dale J. Poirier. Partial observability in bivariate probit models. *Journal of Econometrics*, 12(2):209–217, February 1980.

Maria Ponomareva and Elie Tamer. Misspecification in moment inequality models: Back to moment equalities? *Econometrics Journal*. forthcoming.

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

Christian P Robert and George Casella. *Monte Carlo Statistical Methods*. Springer texts in statistics. Springer, New York, 2nd ed edition, 2004.

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

T. Saarimaa and J. Tukiainen. Coalition formation and political decision making. Technical report, Government Institute for Economic Research, Finland, 2010.

Charles M. Tiebout. A pure theory of local expenditures. *The Journal of Political Economy*, 64(5):416–424, October 1956.

Kenneth E Train. Simulation methods for probit and related models based on convenient error partitioning. Technical Report 95-237, UC Berkeley, 1995.

Eric Weese. Political mergers as coalition formation : Evidence from japanese municipal amalgamations. Technical report, Center for Research on Contemporary Economic Systems, Hitotsubashi University, December 2008.

S. Wernicke. Efficient detection of network motifs. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on*, 3(4):347–359, 2006.

A Proofs

A.1 Proof of Theorem 1

Let $\mathscr{S}$ be the set of all potential coalitions. The unique stable partition can be constructed as follows:

0. Let $k = 0$, $\pi^0 = \emptyset$, and $\mathscr{S}^0 = \mathscr{S}$.
1. Find $S^k_{\max}$ such that $\forall S \in \mathscr{S}^k$, $u(S) < u(S^k_{\max})$.
2. Set $\pi^{k+1} = \pi^k \cup \{S^k_{\max}\}$ and $\mathscr{S}^{k+1} = \{S | S \in \mathscr{S}^k, S \cap S^k_{\max} = \emptyset\}$.
3. If $\mathscr{S}^{k+1} \neq \emptyset$, repeat from 1.

A.2 Proof of Theorem 2

(existence). By construction, Π^* is internally stable. Now take some $\pi \notin \Pi^*$. Then $\exists \{\pi_1, \ldots, \pi_m\} \subset \Pi$ such that $\pi \to \pi_1 \to \cdots \to \pi_m$ and either $\pi_m \in \Pi^*$ or there is a cycle with $\pi_m = \pi_l$ for some $l < m$. If there is such a cycle, then it must contain both mergers and dissolutions. Suppose that $\pi_k \nearrow_S \pi_{k+1}$, and let $S_1^+ \subset S$ be the set of agents that strictly prefer π_{k+1} to π_k. If $\pi_{k+1} \nearrow \pi_{k+2}$ then $S_2^+ \supset S_1^+$ since no agent can be made worse off by a merger. If $\pi_{k+1} \searrow_{S'} \pi_{k+2}$ then $S_2^+ = (S_1^+ \setminus R) \cup P$ 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 $S^+_{m-l+1} = \emptyset$, 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 some subset of S_2^+ cannot be made worse off, and thus S^+ can never be empty. Thus a cycle cannot exist, and $\pi_m \in \Pi^*$. □

(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 $\Pi^{**} = \Pi^*$.[40] □

(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. □

[40] This does not imply that $|\Pi^*| = 1$.

B Data

Population data comes from the 1995 national census, which provides data at the kilometer grid square level. Taxable income per capita is used as a proxy for income per capita, with 1996 tax year data taken from the Asahi Shimbun *Minryoku*. The list of mergers that actually occurred is from the Japan Geographic Data Center. To construct the set of possible coalitions information on which municipalities share a border is derived from shape files for 1995 boundaries from ESRI Japan. These files are also used in conjunction with the census grid square data to calculate the location of generalized median voters.[41]

Municipal financial data is from the '96-'97 and (for comparison) the '06-'07 fiscal year *Shichōsonbetsu Kessan Jōkyō Shirabe*, an official national government report of municipal finances.[42] Because of the large transfers from the national treasury to local governments, this data is handled quite carefully by officials in the central ministries and is generally regarded as accurate, particularly the sections produced by the central government itself. The isolated incidents of fraud reported generally relate to variables reported by the municipalities, which are not used in this paper.

Table 3 gives summary statistics for this data. Units for population and surface area are chosen to be close to the Ministry of Internal Affairs' "reference municipality". There are no missing values in any of the financial, population, or surface area data. In the income data, four values are missing because one merger took place after the data was collected but before it was published, and the old municipalities were not reported. This variable, however, is only used in the preliminary financial regressions described below, and thus the data used in the main estimator does not have any missing values.[43]

[41] A precise calculation of the generalized median is computationally intensive, and infeasible given the number of such values that need to be computed. A simple calculation of median latitude and median longitude, however, yields a good approximation, and is used throughout the paper. Using the mean voter, which is easy to calculate but not consistent with the model presented, yields similar results.

[42] This report also provides by municipality population data from the 1995 census, and surface area data from a 1996 Geographical Survey Institute survey.

[43] The 23 "special wards", which cover roughly the area of pre-war Tokyo City, have powers similar to cities but are excluded from the analysis because any enlargement of this *sui generis* area would likely involve adding more wards, rather than changing the borders of existing ones. The twelve "designated cities", which have some powers normally reserved for prefectures, are omitted from the preliminary financial calculations because their additional responsibilities increase their required spending, but they are included in the rest of the analysis as regular cities. The categories of "core

The initial laws implementing the merger incentive scheme were passed in 1995, and thus it would be optimal to use data from before this point. However, electronic data availability improves around 1995, and most of the data used is from the '95-'97 period. Endogeneity problems seem unlikely, as most of the merger negotiations and approvals occurred in the middle of the 1999-2010 window. In fact, the "Trinity" tax reforms were not even finalized until 2002, and this uncertainty provided municipalities with an incentive to wait until after 2002 to conduct any mergers. Consequently, using data from '95-'97 should not cause difficulties. In particular, the financial data used is calculations by the national government, not actual spending by municipalities, and thus is not vulnerable to "last minute" capital spending, where a municipality planning on merging in 1999 might increase expenditures in its '96-'97 budget.

	Units	Mean	SD
$\tilde{c}_m$ (std. fiscal need, ¥)	100,000,000	54.64	88.87
POPULATION	100,000	0.31	0.65
AREA (sq. km.)	100	1.13	1.33
INCOME (per capita, ¥)	100,000	30.99	4.45
INCOME.INEQ	coef. of var.	76.42	19.47
CITY (vs. town/vill.)		0.20	0.40

Table 3: Summary Statistics

city" and "special city" were created during the merger period, and therefore do not appear in the data used.

C Japanese Local Public Finance

Post-war Japanese fiscal policy emphasized equalization between municipalities, and established national standards for local government services.[44] To ensure that every municipality had sufficient funds to offer services above this minimum level, the national government developed a complicated system of transfers, called the "Local Allocation Tax".[45] The estimated cost $\tilde{c}_m$ for the provision of public services varies from municipality to municipality, based on a formula developed by bureaucrats at the Ministry of Internal Affairs and Communications. The 2009 version of the exposition of this formula (the *Chihō Kōfuzei Seido Kaisetsu*) consists of 600 pages of Japanese legal text, 460 pages of formulae, and 240 pages of reference values; however, as Figure 2 shows, $\tilde{c}_m$ can be approximated by the non-linear regression

$$\tilde{c}_m(P_m) = (\gamma_0 + \gamma_1 P_m^{\theta})v_m, \tag{38}$$

where v_m is a lognormal error term.[46] The fixed cost is thus γ_0, and the average variable cost is $\gamma_1 P_m^{\theta-1}$. The values in Table 4 reflect the fact that the central ministries believed that there were economies of scale in the production of public goods, and thus per capita costs would be higher in municipalities with lower population: the fixed cost estimate is about ¥600 million for Columns II-V, and the marginal cost ($\gamma_1 \theta P_m^{\theta-1}$) is about ¥200,000 per capita at a population of 1,000, but ¥100,000 per

[44] One possible reason for this policy would be to correct a Flatters, Henderson, and Mieszkowski [1974] "fiscal externality"; however, this requires the possibility of migration between municipalities, which is not included in this paper. Dahlby and Wilson [1994] provide an alternative explanation that is consistent with the model in this paper, where the goal of the transfer scheme is equalizing the marginal cost of public funds across municipalities.

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

[46] The underprediction for large municipalities in Figure 2 is not necessarily evidence that Equation 38 is misspecified. Although "designated cities" have been removed, large cities that do not quite meet the criteria to be so designated appear to have been allocated some additional responsibilities on a more *ad hoc* basis. Because the underlying data that the Standard Fiscal Need calculations are based on is not available, the amount allocated for these additional responsibilities is difficult to remove; however, Hayashi et al. [2010] show that the functional form used here matches cost estimates generated by the *dankai hosei* adjustment coefficient, which is the adjustment that creates most of the per capita cost difference between municipalities with large and small populations. Weese [2008] used a simpler specification which matched the data in Figure 2 better for large municipalities: this specification was changed to that in Equation 38 after consulting with the Ministry of Internal Affairs officials responsible for the cost estimates.

	I	II	III	IV	V
(Intercept)	9.57*	5.67*	6.15*	5.94*	6.16*
	(0.19)	(0.12)	(0.17)	(0.12)	(0.17)
θ	0.90*	0.90*	0.87*	0.93*	0.89*
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
POPULATION$^\theta$	137.19*	139.17*	143.01*	146.76*	145.21*
	(2.89)	(3.04)	(3.14)	(3.02)	(3.01)
AREA		3.19*	2.73*	3.13*	2.58*
		(0.10)	(0.10)	(0.10)	(0.09)
INCOME		0.01	0.09*	−0.03	−0.00
		(0.02)	(0.03)	(0.03)	(0.03)
INCOME.INEQ		−0.07*	−0.07*	−0.03*	−0.04*
		(0.00)	(0.01)	(0.01)	(0.01)
POP$^\theta$*INCOME				1.02*	2.11*
				(0.25)	(0.27)
POP$^\theta$*INCOME.INEQ				−0.60*	−0.81*
				(0.07)	(0.07)
PREFECTURE	N	N	Y	N	Y
CITY	N	N	Y	N	Y
N	3220	3216	3216	3216	3216

Table 4: Dependent variable is Standard Fiscal Need ('96-'97)

capita at a population of 100,000.[47] Taking into account both the fixed cost and the variable cost, the per capita cost of providing national minimum level public goods in Ashiyasu village (population 567) would be roughly ¥1.5 million, compared with roughly ¥100,000 in Sakai city (population 790,000).

With almost half of pre-merger municipalities having a population less than 10,000, the decision to provide additional subsidies to smaller municipalities due to their size was an expensive one. To attempt to reduce this inefficiency, in the 1990s the government implemented three policies designed to promote mergers.

First, the government substantially reduced transfers to the smallest municipalities

[47]Column I is provided for reference, but as there are explicit expense categories dealing with land administration (particularly forests and farmland), surface area is an obvious variable to include. The results in Column I show that the null hypothesis that there are no economies of scope between administering land and providing services to residents cannot be rejected: the difference in estimates for γ_1 and θ are both statistically insignificant between Column I and Column II. Comparisons with Columns III-V do yield some statistically significant differences, but using parameter values from these columns does not change the results in the following sections of the paper.

	96-97	06-07
(Intercept)	6.05*	3.62*
	(0.23)	(0.23)
θ	0.89*	0.87*
	(0.01)	(0.01)
POPULATION$^\theta$	133.83*	133.30*
	(5.31)	(8.55)
AREA	3.45*	2.99*
	(0.19)	(0.24)
N	1129	1129

Table 5: Dependent variable is Standard Fiscal Need

by revising the Local Allocation Tax. This can be approximated as

$$T_m^{\text{new}} = \max(\tilde{c}_m^{\text{new}}(P_m) - .75\bar{\tau}Y_m, 0) \tag{39}$$

$$\tilde{c}_m^{\text{new}}(P_m) = (\gamma_0^{\text{new}} + \gamma_1 P_m^\theta)v_m \tag{40}$$

as shown in Table 5, with γ_0^{new} being about 40% lower than γ_0.[48] The government documents describing $\tilde{c}_m^{\text{new}}$ make it obvious that this is not a re-estimate of costs required to provide services, but rather an arbitrary decrease in transfers to small municipalities: to compute a transfer based on $\tilde{c}_m^{\text{new}}$, a set of per capita ceilings are applied to the original components of $\tilde{c}_m$ and the result summed.

Second, municipalities that merged between 1999 and 2006 would not have their transfers lowered due to the merger for at least ten years starting from the date of the merger. That is,

$$T_S^{\text{new}} = \sum_{m \in S} T_m^{\text{new}} \tag{41}$$

would be provided for the decade following the merger.[49] This incentive began to

[48]The sample in Table 5 is restricted to those municipalities that did not participate in a merger in order to have the same sample in both periods. Thus, the change in coefficients represents a change in national government transfer policy on the same group of municipalities during the period in question. Inflation during this period was negligible. The difference in the POPULATION coefficient is statistically insignificant, and the difference in θ and the AREA coefficient is marginally statistically significant (p roughly 0.05). The only economically significant change is in the intercept, indicating that the new policy effectively involved a lump sum reduction compared to the old policy.

[49]An intermediate amount between Equations 39 and 41 was offered for years 11-15 following a merger.

be phased out in 2006, which motivated many municipalities to finalize mergers in 2005. By 2006 there were only 1,844 municipalities remaining, down from 3,255 at the start of the merger period. A small number of mergers occurred during the phase-out period, reducing the final number of municipalities to 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.[50]

A final incentive for mergers was the *Gappei Tokureisai*, special subsidized bond issues allowed for municipalities planning amalgamation.[51] The value of these bonds 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.

[50]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 may be an interesting area for further work.

[51]The official explanation for these bonds was to eliminate any direct financial cost of merging, such as the construction of a new city hall. The merger 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 very small effect on incentives to merge, and thus for simplicity the the bonds are treated entirely as an additional incentive, with the direct financial costs of merging ignored.

D Random Draws of Partitions via Markov Chains

Let the state space $\mathcal{X}$ be the set of all partitions, and the transition matrix P be defined as follows:

$P_{xy} = k$ if y can be created by either breaking apart one coalition in x into subcoalitions, or merging coalitions in x together to create one new coalition,

$P_{xy} = 0 \quad \forall y \neq x$ that do not meet the above condition,

$P_{xx} = 1 - \sum_{y \neq x} P_{xy}$.

Since the state space is finite, this creates a valid transition matrix for sufficiently small k. P describes a reversible Markov chain, since a transition from x to y via merging implies a possible transition from y to x via a breakup, and vice versa. The chain is connected, since any partition can be obtained by first breaking all coalitions down to singletons, and then constructing the desired partition. There is thus a unique stationary distribution, since the chain satisfies the detailed balance condition [Robert and Casella, 2004]. Moreover, the stationary distribution gives equal probability to each state, and thus draws from the stationary distribution are equivalent to random draws from the set of all partitions. These draws can be performed using the Metropolis-Hastings algorithm.

Unfortunately, the number of transitions that need to be considered is too large to be computationally feasible. The number of ways a size fifteen coalition could be broken down into sub coalitions could be as high as a billion, and thus it is not practical to enumerate all the possible transitions. Instead, the approach introduced by Wernicke [2006] for the RAND-ESU algorithm will be used. Let p be a length fifteen vector of "cut probabilities" and let Q be a set of subcoalitions of S that form a partition of S. Then enumerate a member of Q with probability $\prod_{i \in \{1,\ldots,|S|\}} p_i$. This gives each member of Q an equal probability of being enumerated, but allows for a much smaller, randomly selected set to be considered. So long as this randomly selected $\tilde{Q}$ is re-randomized each time a given state is reached, then the properties of the above Markov chain are unchanged.