Who Merges with Whom? An Econometric Analysis of Local Government Mergers[†]

Tuukka Saarimaa

VATT
P.O. Box 1279 (Arkadiankatu 7), FI-00101 Helsinki, FINLAND.
E-mail: tuukka.saarimaa@vatt.fi

and

Janne Tukiainen
VATT and HECER
P.O. Box 1279 (Arkadiankatu 7), FI-00101 Helsinki, FINLAND.
E-mail: janne.tukiainen@vatt.fi

Very preliminary draft, this version March 2010
Comments welcome, please do not quote without authors' permission

Abstract
In 2005 the Finnish government initiated a plan to restructure municipal service production by encouraging municipality mergers. We analyse three questions concerning municipal mergers. First, we study the effectiveness of the central government's merger subsidy scheme. Second, we study the association of local politics with merger decisions. Third, we study how municipal population and economic characteristics are associated with mergers. The most robust results from our coalition level logit analysis are that existing cooperation has a strong positive effect on the likelihood of merging and that distance prohibits merging. Population of the coalition also seems to matter so that larger coalitions in terms of population are more likely but only up to a point. On the other hand, using a partial observability bivariate probit model with population sorting, we find that the merger subsidy has a surprising heterogeneous impact on the merger activity. Another finding is that local political factors seem to be relevant for merger decision-making.

Keywords: Municipality mergers, bivariate probit, two-sided decision-making
JEL codes: H71, H72, H77, C35

[†] We thank Elias Einiö, Ari Hyytinen, Aki Kangasharju, Jani-Petri Laamanen, Antti Moisio, Jaakko Pehkonen, Matti Sarvimäki and Otto Toivanen for helpful comments. Niklas Jahnsson provided excellent research assistance. We also thank seminar participants at FEAAM-Tampere and VATT for useful discussions.

1 Introduction

Local governments play an important role in the economies of many countries. This is especially true in the Nordic countries like Finland where there has been a long tradition of decentralized political decision-making. In Finland, public services are provided by two tiers of government where municipalities constitute the local level and municipal co-operation an implicit third level. Because of the variety of tasks assigned to them, municipalities are of considerable importance to the whole economy. The GDP share of municipality spending is roughly 18 percent and they employ around 20 percent of the total workforce.

In this paper, we analyse empirically a recent surge of municipality mergers in Finland.[1] We study three important questions concerning the mergers. First, we study the effectiveness of a central government merger subsidy scheme. The central government grants merger subsidies to merging municipalities according to a complex step function of the number of disappearing municipalities and populations of the municipalities involved in the merger. This type of subsidy is in place only from 2008 to 2013. There is large variation in the per capita merger subsidies among the potential mergers which allows for the evaluation of the effects of this exogenous policy on merger activity. Second, we study the association of local politics with merger decisions. We observe some interesting characteristics of the municipal councils that allow us to study whether council members' private incentives are associated with their decisions concerning mergers. We also study how councils' political party compositions are correlated with the merger decisions. That is, we are interested in whether the local political structure hinders otherwise sensible mergers or promotes unwise mergers. Third, we study how municipal population characteristics, characteristics related to costs of providing statutory services and municipal tax bases are associated with merger decisions.

These questions are interesting in their own right, but the answers also serve as a first step in evaluating whether municipality mergers are an effective way of meeting the ultimate goal of the central government, which is to make municipal service production more efficient. They also shed light on the question whether the central government should encourage mergers financially, impose mergers upon municipalities or respect local autonomy and leave the merger decisions up to local councils.

[1] The institutional setting is described in section 2.

Political coalition formation has received considerable attention from both policy makers and economists. For example, Miceli (1993) and Alesina and Spolaore (1997, 2003) emphasize the fundamental trade-off between economies of scale (which favours merging) and regional heterogeneity in preferences (which favours separation). Bolton and Roland (1997) offer similar arguments. Bradford and Oates (1974) argue that internalising inter-jurisdictional spillovers may also justify consolidation of jurisdictions. On the other hand, Ellingsen (1998) argues that economies of scale are a questionable rationale for merging since economies of scale can be just as easily exploited through contracting or cooperation. In fact, cooperation in service production is quite common among Finnish municipalities.[2] Cooperation in producing some services but not others is a way to deal with the fact that different services may have different economies of scale.

One must therefore ask what the additional benefits or costs are that merging provides compared to cooperation. One apparent difference is that merging depends on and alters political power structures in a way that cooperation does not. There are many aspects to this. First, it is obvious that the costs and benefits of mergers are not equally divided among all members. Clearly there are winners and losers. In a democratic system with majority voting, mergers may not occur even when merging would be efficient. It is important to note that when all parties must accept the merger, a majority against the merger in just one municipality is enough to block it. Second, efficient mergers may be blocked or inefficient mergers may take place because of local politicians and bureaucrats self-interest.[3] For example, municipal employees who are also council members may have private incentives either to vote for or against merging regardless of the preferences of their constituents. This is especially true when merging is encouraged mainly on cost efficiency arguments. Third, council members of a particular party may find a merger undesirable because in the new municipality their party's share of council seats would be lower or that they would face stiffer competition from members of their own party. Fourth, large municipalities have a stronger

[2] There are 226 joint authorities set up by two or more municipalities.

[3] Ellingsen's (1998) theoretical model predicts that excessive merging will not take place, but municipalities may forego a welfare enhancing merger. In his median voter model, the departure from optimality is caused by the selfishness of the politically powerful combined with the impossibility of perfect contracting. However, Ellingsen's model is based on maximising social welfare (or majority's welfare) and does not incorporate public choice aspects.

leverage in negotiations between the central government and the municipal sector.[4] In Finland, municipalities have organised an association (the Association of Finnish Local and Regional Authorities or *Kuntaliitto*) that facilitates municipal cooperation but also acts as a strong lobbyist and a counterforce against the central government. Of course, the larger a municipality is the stronger is its influence on central government's policy through the Association.

Another important duty of municipalities is to determine land use zoning policies and organize public transport. These may be more effectively coordinated within a larger regional entity than an individual municipality. Furthermore, merging may create financially stronger municipalities and a reduced need for tax base equalisation by the central government. Moreover, merging may bring more cost savings and lower taxes through more efficient bureaucracy. Larger and fiscally sound municipalities may also be able to attract more skilled workers to the municipality sector than smaller ones.

On the other hand, if households have sorted themselves to municipalities in the spirit of Tiebout (1956) it is not clear whether one should promote or even expect merger activity. Some recent developments in Finland have, however, changed the operating environment of many municipalities. A major recession in the early 1990s hit different regions with a different force. The subsequent recovery was regionally uneven and regional disparities started to grow. In particular, during the last ten years or so, regional disparities have grown rapidly because the workforce has concentrated into few growing centres where employment possibilities are better. This means that both the municipalities that are losing population and the ones gaining have experienced major changes in their operating environment. These changes may be enough to encourage municipalities to search for new ways to cooperate in service production or even to merge.

Econometric modelling of merger decisions is complicated for a number of reasons. First, merger decision-making is two-sided and we observe a merger only if all merger parties agree to merge. Second, every municipality faces multiple potential merging partners. Third, municipal merger choices are spatially interdependent so that a merger between two or more

[4] Finnish municipalities are self-governing bodies by constitution which means that central government cannot assign new responsibilities to municipalities without passing new legislation. However, the central government may, for example, change municipal tax bases and assign limits to property tax rates.

municipalities changes the choice set of adjacent municipalities. Finally, there is the possibility of one-to-many matching, meaning that coalitions that are larger than two are possible. This means that the actual number of potential partners and coalitions is very large.

Recently, Gordon and Knight (2009) and Weese (2009) have developed novel estimation strategies to tackle these estimation problems. However, for various reasons, these methods are not suitable for analyzing our research questions. We solve the problem of one-to-many matching by combining Wernicke's (2005) network detection algorithm with choise-based sampling (case-control). We then analyze coalition formation both by using a simple logit analysis where potential coalitions are used as the unit of observation and also by estimating Poirier's (1980) bivariate probit model with partial observability which can accommodate for the two-sided decision-making.

The rest of the paper is organised as follows. Section 2 presents a more detailed description of Finnish municipalities and the institutional background for municipality mergers. Section 3 looks through previous empirical studies on local government mergers and coalition formation. Section 4 describes the econometric methods and data. Section 5 presents the empirical results and section 6 concludes.

2 Institutional background

Finnish Municipalities are important both in terms of their functions and in terms of their economic significance. Unlike in the mainstream economic models of local public finance, where local governments provide local public goods, the bulk of Finnish municipalities' expenditures come from producing welfare services with a strong redistributive character. Municipalities are responsible for providing the most of social and health care services, primary education and culture services. In addition, municipalities provide the basic environment and technical infrastructure services along with public transportation. Municipalities also have a zoning monopoly within their borders.

Another important feature of the Finnish system is the large number of municipalities with a huge variation in municipal population size. The largest municipality is the country's capital, Helsinki, which has over half a million inhabitants, whereas the smallest mainland municipality, Suomenniemi, has only 801 inhabitants. The median municipal population size

is less than 6,000. The number of municipalities has diminished considerable since the 1940's as can be seen from Figure 1. All the mergers have been voluntary. Finland is also sparsely populated and population density varies a great deal between municipalities.

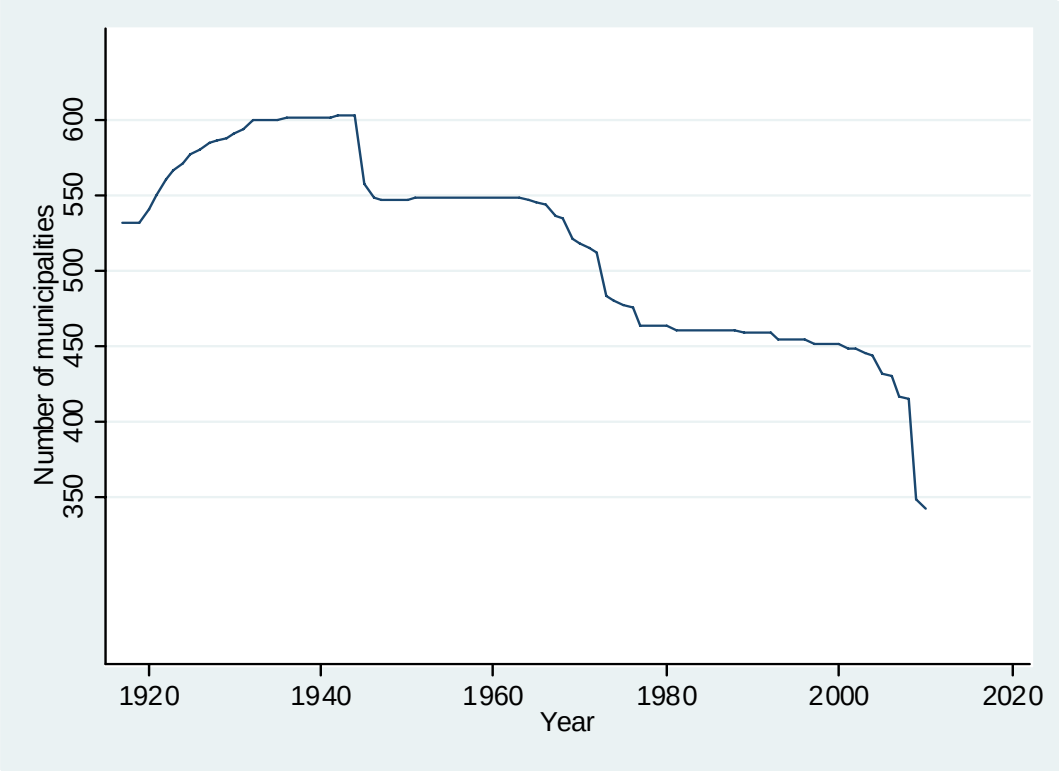

Figure 1. Number of municipalities in Finland, 1917–2010.

Municipalities fund their spending mostly through their own revenue sources. The most important revenue source is the proportional municipal income tax which constitutes 40 percent of the revenues. Municipalities are also entitled to a share of the corporate income tax paid by corporations in the municipality, but corporate taxes make up only about 4 percent of revenues. The significance of the property tax is even smaller at only 2.5 percent of revenues. Due to the fact that most municipal services in Finland are statutory, and the large disparities in municipal population structure and density, a central government grant system is used to equalise local cost and revenue disparities. The grant system compensates municipalities with unfavourable cost conditions, expensive age structure and social problems. The main components of the current grant system are the blocks for health care and social services and for education. The system also includes a minor general grant and a separate system that

equalises tax bases. On average, central government grants cover 20 percent of total municipal finances. The rest of municipal revenues consist of user-fees and sales incomes.

As already mentioned, regional disparities have grown because population has concentrated into a few growing employment centres. In many small and remote municipalities this has led to unsustainable deficits and the grant system has become the most important source of revenue for these municipalities. Currently, the grant system covers more than 50 percent of all revenues for every fourth municipality. Due to an expected increase in municipal spending and disparities in revenue base the central government initiated a plan (so-called PARAS initiative) that aimed at reforming municipal revenue structure and more importantly making the production of statutory municipal services more efficient. The main tool for strengthening the operating environment of municipalities set forth in the new law was municipality mergers. The law clearly states that municipalities should have strong enough revenue and labour force bases to cope with the production of statutory public services.

In order to facilitate mergers, the central government grants merger subsidies to merging municipalities. The subsidy scheme consists of two parts. The basic part is determined according to the population of the new municipality and the combined populations of the non-largest municipalities involved in the merger. The central government also encourages multiple mergers with a supplement part. A supplement amount is paid if the number of municipalities diminishes at least by two and increases as the number of disappearing municipalities rise. This subsidy scheme is in place from 2008 to 2013. The subsidy to a given merger is increased by 80 percent if the merger took place in 2008 or 2009 and by 40 percent if the merger takes place in 2010 or 2011. We concentrate on mergers that took place in 2008 and 2009. These mergers were decided by municipal councils elected in 2004 and thus we can treat these mergers as a result of one period coalition formation game.

Table 1 illustrates the subsidy scheme. It is important to note that, in addition to the subsidy, the government guarantees that grants do not to fall for five years after the merger even if the new municipality would be entitled to a smaller overall grant than the merging municipalities individually. The subsidy scheme is clearly designed to promote economies of scale from the mergers as the amount paid is higher if the new municipality has over 20,000 inhabitants. However, the per capita subsidy is still small for very large municipalities. In addition, the subsidy scheme rewards savings in bureaucracy because the amount paid is the higher the

higher is the number of municipalities involved in the merger. It is important to emphasize that the discontinuities in the subsidy scheme offer us a way to distinguish the effects of the subsidy from the role of economies of scale. In this study, we make a plausible assumption that the subsidy scheme is exogenous. In the future, we plan to use these discontinuities to a greater extend and utilize a regressions discontinuity design to analyze the effects of these mergers.

Table 1. The central government subsidy scheme in 2008–2009.

Population in thousands (first number is the new municipality, second number is the combined of non-largest municipalities):	Number of disappearing municipalities:									
	1	2	3	4	5	6	7	8	9	10
over 20, over 10	7.20	8.46	9.72	10.98	12.24	13.50	14.76	16.02	17.28	18.54
over 20, 5-10	6.48	7.74	9.00	10.26	11.52	12.78	14.04	15.30	16.56	17.82
over 20, less than 5	5.76	7.02	8.28	9.54	10.80	12.06	13.32	14.58	15.84	17.10
under 20, over 7	5.40	6.66	7.92	9.18	10.44	11.70	12.96	14.22	15.48	16.74
under 20, 3.5-7	4.50	5.76	7.02	8.28	9.54	10.80	12.06	13.32	14.58	15.84
under 20, under 3.5	3.60	4.86	6.12	7.38	8.64	9.90	11.16	12.42	13.68	14.94

Notes: Figures in the table are millions of Euros.

3 Previous empirical studies

Political coalition formation has received considerable attention from both theoretical and empirical economists. Alesina et al. (2004) examine the number of political jurisdiction, such as school districts and municipalities, within counties in the U.S. In particular, they test whether a trade-off between economies of scale and population heterogeneity can explain the number and size of jurisdictions. Using data from the period of 1960–1990, Alesina et al. find evidence of a trade-off between economies of scale and heterogeneity in race and income. However, they do not analyse actual merging decisions of these entities.

Brasington (1999) uses Poirier's partial observability bivariate probit which allows for two-sided decision-making to study school district consolidations in Ohio U.S.[5] He finds that racial composition and the difference in racial composition have no independent effect on consolidation, nor do income levels or school quality. He argues that consolidation happens

[5] Tiller and Jakus (2005) also apply Poirier's model when they explain cooperation in municipal solid waste management in Tennessee counties.

mainly with cost savings in mind. In particular, he finds a general tendency of small and large school districts to merge, whereas middle-sized districts merge less often.

Brasington (2003a) uses the same data as Brasington (1999) but arranges the data according to population in order to see whether small and large communities differ in their preferences to consolidate. Brasington (2003a) also finds clear differences in the way big and small communities react to differences in population size. The larger the population gap between the big community and the small community, the more the big community wants to consolidate. The opposite is true for the small community. This means that a big community gains scale economies while maintaining political control over schooling decisions. This result is in line with predictions from Ellingsen's (1998) model.

Brasington (2003b) uses the same sorting idea as Brasington (2003a) but now sorting is done according to income and race (percentage of white). That is, richer communities are compared to poorer and whiter to darker in each potential consolidation pair. Brasington (2003b) finds that the higher property values in a community the less likely it is to consolidate. This is true for richer and poorer potential partners. When data are arranged according to percentage white property values effect is still negative but it is statistically significant only for less white communities.

The main problem in Brasington's (1999, 2003a, 2003b) models is that the correlation coefficient of the two probit equations is 0.99 or 1 in many of his models indicating a corner solution to the maximization problem. According to Butler (1996), error correlation that approaches -1 or 1 violates the regularity conditions of maximum likelihood, and causes the Hessian to be noninvertible or implausible *t*-values. A further problem with all of Brasington's papers is that he uses post-merger district characteristics to explain mergers. If merged school districts tend to become more similar after merging, using post-merger data from 1990 to explain mergers in the 1960's is bound to produce errors in analysing the effects of heterogeneity. We use the estimation strategy as Brasington but address both of these problems properly.

Sørensen (2006) regresses Norwegian municipal leaders' (mayors, council members and chief administrators) stated preferences on consolidations on municipal characteristics and potential efficiency gains from merging. Based on a questionnaire and municipal characteristics he

finds that these groups are more favourable to merging when potential mergers offer large efficiency gains. On the other hand, revenue disparities seem to be significant impediments to consolidations in Norway. However, Sørensen (2006) is unable to match the potential merging partners precisely, but instead analyses individual municipality leaders' preferences.

More recently, Gordon and Knight (2009) and Weese (2009) have proposed methods to analyse spatial mergers that take into account also other potential complications. The spatial simulated moments estimator by Gordon and Knight (2009) takes into account two-sided decision-making, spatial interdependence and multiple potential partners but does not allow for one-to-many matching. Gordon and Knight (2009) apply their estimator to school district mergers in Iowa in the 1990's. They find that state subsidies were an important factor influencing merger decisions of school districts. They also find that economies of scale matter. Furthermore, heterogeneity in education level of district population (percentage with a college degree), spending and distance between potential merger partners makes merging less likely. Unfortunately, due to a computational burden, their estimator allows only a handful of variables at a time.

Weese (2009) introduces a method of estimating the structural parameters of a political coalition formation game and applies the method on municipality mergers in Japan. Weese's estimator is novel because it allows also for one-to-many matching. This is accomplished by clever application of Wernicke's (2005) algorithm. However, the underlying Alesina and Spolaore (1997) type structural model in Weese (2009) does not account for our questions of interest like the role of local politics. He estimates a coalition formation game using coalition level characteristics. This set up does not easily allow for the individual political characteristics of different potential merger partners. Moreover, the full use of our rich data would be computationally challenging also in Weese's framework. Because the estimator is computationally intensive Weese uses only a handful of arguments in the coalition utility function. He finds, however, that high government services and low population are preferred. His results are not clear on whether there was a strong preference for status quo in the pre-merger period or not.

It is clear that a number of important factors contributing to merger decisions are left out in both Weese's and Gordon's and Knight's empirical applications of their estimator. Since our main contribution is to use a rich data set with many political variables of interest, these

computationally intensive methods are not very useful for our purposes. Moreover, the underlying theoretical models in these studies do not allow for our public choice questions. Neither Gordon-Knight nor Weese estimator is suited for analysing the effects of local political conditions or politicians self-interest on merger decisions, since they are based on maximising utility functions of coalitions or individual jurisdictions (or populations of coalitions) but not of the political decision makers'. For these two reasons, the characteristics of all potentially merging municipalities' councils cannot be easily included in their analysis.

It should also be mentioned that Brasington (1999, 2003a, 2003b) and Gordon and Knight (2009) analyse U.S. school districts that provide elementary schooling and are financed mainly through property taxes. Our analysis is more complex because Finnish municipalities provide a wide range of public services.

4 Econometric model and data setup

Econometric modelling of merger decisions is complicated for a number of reasons. First, merger decision-making is two-sided and we observe a merger only if all merger parties agree to merge. Second, every municipality faces multiple potential merging partners. Third, municipal merger choices are spatially interdependent so that a merger between two or more municipalities changes the choice set of adjacent municipalities. Finally, there is the possibility of one-to-many matching, meaning that coalitions that are larger than two are possible. This means that the actual number of potential partners and coalitions is very large.

There are three different estimators in the literature that account for some or all of these features. First, Poirier's (1980) bivariate probit model with partial observability can accommodate for two-sided decision-making but not the other important features. This approach has been used by Brasington (1999, 2003a and 2003b). Second, the matching estimator proposed by Gordon and Knight (2009) takes into account two-sided decision-making, spatial interdepence and multiple potential partners but does not allow for one-to-many matching. Merges including multiple municipalities are a prominent feature of our data which makes the Gordon-Knight estimator infeasible in our case. Third, Weese (2009) introduces a method of estimating the structural parameters of a political coalition formation game. Weese's estimator is novel because it allows for all the above features including one-to-many matching.

Since our main contribution is a rich set of interesting explanatory variables, we will not estimate a complex structural model. In other words, we are facing a trade-off between using our data to the full extent in answering important questions and being able to deal with relevant econometric problems. We choose to favour the first to the second benefit. Therefore, we use two reduced form approaches. First, we analyse coalition formation using a simple logit analysis in a choice-based sampling framework (case-control) where the coalition is used as an observation unit. The sampling procedure is as follows. For each merger of size n (n = number of merging municipalities), we sample 4 potential mergers of equal size that did not actually take place as controls. We repeat this for all n that actually took place (2, 3, 4, 5, 6 and 10). The sampling from all potential coalitions can be done using FANMOD software by Wernicke and Rasche (2006). The FANMOD software takes into account municipality borders so that municipalities can form a coalition only if all of them share a geographic border. Naturally, some municipalities may merge even if they are not neighbours. This happens when they are a part of a coalition including multiple municipalities that jointly form a geographically coherent municipality.

One problem with this procedure is that large coalitions may take many different forms and some of them may be quite unrealistic, and thus, would not serve as very good controls.[6] For example, some coalitions of size 6 may be string-like and result in high transportation costs. In order to circumvent this problem, we restrict the potential coalitions so that they cannot cross county lines. Figure 2A in the appendix shows municipality and county borders. From the figure it's clear that coalitions that stay within county borders (bold line) are much less likely to be problematic in this sense. This choice is also supported by the fact that no mergers took place across county lines. There are at least two reasons for this. First, counties constitute a middle-level in regional policy-making in Finland even though counties have very limited political power. For example, all regional administrative authorities by the central government should follow county division. This makes counties a natural co-operation environment for municipalities even in the absence of formal mergers and is probably the main reason, along with cultural identity and such, why merging across a county line does not happen. Second,

[6] Weese (2009) faced a similar problem because the largest merger that took place during his analysis period in Japan included 15 municipalities. This creates problems for Weese's estimator because the number of potential coalitions with a size of fifteen or less is so huge. Weese circumvents the problem by introducing an upper bound the number of bordering municipalities that any potential coalition can have. This upper bound basically restricts the geographic shape of the coalitions to be reasonable. In our case restricting the mergers to within county mergers only takes care of this problem.

county borders are almost identical to hospital districts and every municipality has to belong to a hospital district.[7]

The logit analysis reveals us which types of coalitions are likely to form. However, we are also interested on how individual municipalities' own and potential merger partners' characteristics and councils are associated with mergers. In answering these questions we follow the approach by Brasington (2003a, 2003b) and use the Poirier model. Although this model is not perfect in our setup it accounts for the most important feature of our data, which is two-sided decision-making where each potential merger partner has a veto power on the merger. Brasington's (2003a and 2003b) insight in using the Poirier model was to sort the data in a way that gives meaningful interpretations of the coefficients in the bivariate probit equations. We follow Brasington (2003a) and sort the data according to municipality's population. This allows us to concentrate on the political power structure which changes differently for smaller and larger merging partners.

Let y_{1i} and y_{2j} denote the merger decisions of larger municipality i (type 1) and its smaller potential merger partner j (type 2). Poirier's (1980) bivariate probit model can be written as

$$y_{1i} = 1\left(\mathbf{x}'_{1j}\boldsymbol{\beta}_1 + u_{1i} > 0\right), \tag{1}$$

$$y_{2j} = 1\left(\mathbf{x}'_{2j}\boldsymbol{\beta}_2 + u_{2j} > 0\right), \tag{2}$$

where 1(.) is an indicator function that equals one if the statement in the parentheses is true and zero otherwise. The error terms follow a standardised bivariate normal distribution with correlation ρ. Thus, the model allows for same unobservables affecting the merger decision of both parties. Poirier's model differs from the standard bivariate probit in that instead of observing both y_1 and y_2 we only observe their product $z = y_1 * y_2$.

Table 2 clarifies the data setup. Panel A illustrates potential mergers with only two partners. The first column in Table 2 is simply the identification number. The first municipality in each potential match pair is the larger one corresponding to equation (1). This sorting of the data

[7] We also extract the counties of Kainuu and Lapland from our analysis. Kainuu county is experimenting with a level council, and thus, the municipalities will not merge while the experiment is ongoing. Lapland is an outlier in the data, with large and low population municipalities. Drawing control coalitions from Lapland would not be a good idea. There were no mergers in Lapland during our analysis period.

allows one to distinguish the effects of explanatory variables separately for “large” and “small” municipalities, an important insight made by Brasington (2003a and 2003b). One should note, though, that a given municipality may be either “large” or “small” depending on the potential match pair.

Table 2. Example of data setup.

Panel A:								
Potential match pair	Mun. 1	Mun. 2	Merged	Population 1	Population 2	Difference in population 1	Difference in population 2	Gov. Subsidy
1	A	B	0	20 000	9 000	11 000	-11 000	400
2	D	A	1	67 000	20 000	47 000	-47 000	200
3	C	B	0	21 000	9 000	12 000	-12 000	1500
6	G	H	0	14 500	5 000	9 500	-9 500	320
Panel B:								
Potential match pair	Mun. 1	Mun. 2	Merged	Population 1	Population 2	Difference in population 1	Difference in population 2	Gov. Subsidy
..	..	..	..	..	..			..
101	I	J,K,L	1	60 000	42 000	18 000	-18 000	900
102	I,K,L	J	1	81 000	21 000	60 000	-60 000	900
103	I,J,L	K	1	90 000	12 000	78 000	-78 000	900
104	I,J,K	L	1	93 000	9 000	84 000	-84 000	900
..	..	..	..	..	..			..

A particular challenge using the Poirier model comes from mergers that included more than two municipalities. In these cases, we model the situation so that each municipality chooses individually whether to join the coalition. Panel B in Table 2 illustrates. Municipalities I, J, K and L have merged. We treat this observation so that in each case, one of the municipalities considers merging with the three others that form a coalition. The down-side of this procedure is that now the coalition formed by the three other municipalities is treated as if it really was a decision-making entity. Clearly this is not the case in reality. However, this is the best option we have come up with so far. All the potential coalitions from our FANMOD sampling procedure are treated the same way.

We have linked data on municipal mergers from 2008–2009 to municipal characteristics obtained from Statistics Finland and to characteristics of municipal council members obtained from the Ministry of Justice and the Local Government Pensions Institution. The councils were elected in 2004 for a four year term and they made the decision regarding mergers in 2008–2009. Any future mergers are decided by councils elected in 2008. The 2009 municipal

division was used already in the 2008 elections. Restricting the analysis to a single council and a single subsidy scheme, we can think of this set up as a one period coalition formation game.

The starting point for our empirical specification are the facts that through merging municipalities probably look for economies of scale but face a trade-off in terms of matching service production to heterogeneous preferences of a larger municipality. On the other hand, large municipalities may even face diseconomies of scale. To examine both possibilities we follow Gordon and Knight (2009) and use a fairly general average cost function. From the perspective of municipality *i* the efficiency gain or loss from merging with municipality *j* can be expressed as

$$\ln\left[\frac{c(N_i)}{c(N_i+N_j)}\right].$$

where c(.) is the average cost function for public goods and services. Specifying the function as $c(N)=N^{\beta+\gamma N}$ yields the following measure for efficiency gains

$$\ln\left[\frac{c(N_i)}{c(N_i+N_j)}\right]=\beta\underbrace{\left[\ln(N_i)-\ln(N_i+N_j)\right]}_{\text{economies of scale}}+\gamma\underbrace{\left[N_i\ln(N_i)-(N_i+N_j)\ln(N_i+N_j)\right]}_{\text{diseconomies of scale}}.$$

Now β can be interpreted as an estimate of the role of economies of scale and γ as an estimate for diseconomies of scale.[8] This measure is used in the Poirier analysis only. In addition to economies of scale measures, we include other municipal characteristics, measures of heterogeneity of municipalities in coalitions and characteristics of municipal councils.

Some descriptive statistics of our coalition level data are presented in Table 3. The actual mergers and the randomly sampled control coalitions are surprisingly similar. Even their council characteristics are quite similar. However, the actual mergers are smaller in terms of

[8] See Gordon and Knight (2009) for details.

population. Actual mergers are also clearly more likely to have co-operated already before the merger. Existing cooperation here means that the municipalities in the coalition had a joint authority in producing basic health case services. The maximum distance from other municipalities to the largest municipality is larger in the control coalitions (23 kilometres vs. 34 kilometres.).

Table 3. Descriptive statistics for coalitions.

	Merger = 1		Merger = 0	
	Mean	Std. dev.	Mean	Std. dev.
Number of observations	33		132	
population (10,000)	3.44	3.30	5.13	9.04
mean of taxable income (€10,000)	1.23	0.17	1.16	0.18
dependency ratio	0.36	0.03	0.36	0.03
existing cooperation	0.52	0.51	0.16	0.37
language	0.06	0.24	0.10	0.30
merger subsidy (€1,000 per capita)	0.33	0.22	0.37	0.28
maximum distance to largest municipality (100 km)	0.23	0.11	0.34	0.20
total land area	0.11	0.06	0.15	0.19
central government grants (€10,000 per capita)	0.13	0.05	0.15	0.05
deficit[a]	0.21	0.48	0.09	0.29
std.dev. population	1.63	2.34	2.16	4.46
std.dev. taxable income	0.17	0.12	0.14	0.09
std.dev. tax rate	0.01	0.00	0.004	0.003
sdt.dev. population density	0.37	0.34	0.32	0.28
std.dev. depency ratio	0.02	0.02	0.03	0.02
reduction in council size	0.45	0.14	0.45	0.14
Herfindahl index	0.31	0.12	0.31	0.09
same largest party	0.73	0.45	0.71	0.45
left-wing party share	0.31	0.12	0.28	0.09
municipal employees' share in council[b]	0.08	0.03	0.08	0.03
municipal employees' share of population	0.06	0.01	0.06	0.01

[a] *The number of municipalities in the coalition that had four year running deficit smaller the €1,000 per capita. This is a limit set by the ministry of finance. If a municipality is below this limit it has to give the ministry of finance a briefing on how to deal with the deficit.* [b] *Council members who are employed by the municipality they live in.*

Table 4 presents descriptive statistics for municipal level data sorted according to population size as explained earlier. As expected, the larger merging partners are clearly larger in terms of population. But other than that the smaller and larger merger partners are strikingly similar in terms of their characteristics. The smaller municipalities are slightly more concentrated politically as indicated by the Herfindahl index but the difference is small.

Table 4. Descriptive statistics for municipalities sorted according to size.

	Smaller		Larger	
	Mean	**Std. dev.**	**Mean**	**Std. dev.**
Number of observations	420		420	
population (10,000)	0.75	1.37	5.12	6.70
mean of taxable income (€10,000)	1.09	0.18	1.16	0.17
dependency ratio	0.38	0.02	0.36	0.03
cooperation	0.16	0.37	0.16	0.37
merger subsidy (€1,000 per capita)	0.34	0.24	0.34	0.24
distance to largest municipality (100 km)	0.24	0.20	0.24	0.20
total land area	0.20	0.22	0.20	0.22
deficit (€1,000 per capita)	0.22	0.88	0.41	0.73
difference in population	4.37	6.15	4.37	6.15
difference in taxable income	0.14	0.11	0.14	0.11
difference in tax rates	0.005	0.004	0.005	0.004
reduction in council size	0.55	0.15	0.55	0.15
Herfindahl index	0.39	0.15	0.31	0.10
same largest party	0.71	0.45	0.71	0.45
left-wing party share	0.25	0.13	0.29	0.10
municipal employees' share in council[b]	0.08	0.05	0.08	0.03
municipal employees' share of population	0.06	0.02	0.06	0.01

[a] *A four-year running deficit which can be either negative or positive.*

[b] *Council members who are employed by the municipality they live in.*

5 Results

The results from the coalition level logit models are presented in Table 5. The logit coefficients are consistent even-though the sampling is choice-based. The table includes three models. In the first model, we mainly include variables that describe the characteristics of the potential coalitions in terms of population, tax base, financial conditions and cost structure. The second model includes heterogeneity measures of the municipalities involved in potential coalitions. In the third model, we add council characteristic in order to capture the effects of local politics.

The most robust results are that existing cooperation through a joint authority in producing basic health care services has a strong positive effect on the likelihood of merging and that distance prohibits merging. Population of the coalition also seems to matter so that larger coalitions in terms of population are more likely but only up to a point. The optimal municipality population size implied by the logit results is around 60,000. The central government's merger subsidy enters with a negative sign but the coefficient is clearly not

significant. Interestingly, the more variation there is in the tax rates of the municipalities in a potential coalition the more likely a merger is.

The coefficient for the Herfindahl index which measures political fragmentation in councils is positive and highly significant. This indicates that the more concentrated the political power is at coalition level the more likely the coalition is to form a municipality. Furthermore, the share of left-wing parties is positively associated with merging.

Table 5. Results from coalition level logit models.

	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
constant	-2.47	10.00	-29.24*	14.08	-32.61**	17.98
population	0.27	0.33	1.28**	0.50	1.37*	0.71
(population)^2	-0.03	0.02	-0.08**	0.03	-0.12**	0.04
mean of taxable income	6.55	4.08	12.62*	5.58	13.83	7.75
dependency ratio	-29.18	23.86	6.60	27.80	-41.56	42.00
cooperation	1.90**	0.57	2.98**	0.86	4.24**	1.33
language	-0.97	1.26	-1.02	1.54	0.53	2.07
merger subsidy	-1.84	1.79	0.40	1.91	-2.89	2.94
number of municipalilties	0.48*	0.22	0.57*	0.29	0.32	0.52
maximum distance	-7.41**	2.64	-14.03**	4.37	-24.54**	7.16
total land area	-1.10	2.72	0.29	3.01	1.95	8.41
central government grants	29.42	15.94	40.61*	19.36	57.38	31.05
deficit	1.66*	0.71	2.23*	0.93	3.51*	1.43
std.dev. population			0.26	0.27	0.93*	0.44
std.dev. taxable income			-1.38	4.85	1.07	6.28
std.dev. tax rate			266.9*	104.6	515.0*	173.3
sdt.dev. population density			4.88**	1.58	6.63**	2.34
std.dev. depency ratio			-20.18	25.63	-36.93	32.64
reduction in council size					16.64*	6.78
Herfindahl index					25.59**	8.83
same largest party					-0.16	0.86
left-wing party share					10.48*	5.29
municipal employees' share in council[a]					25.32	14.60
municipal employees' share of population					-35.76	34.47
Number of obs.	165		165		165	
Log-L	-56.13		-43.58		-32.52	
Pseudo R^2	0.32		0.47		0.61	

*Notes: ** and * indicate statistical significance at 1and 5 percent level, respectively. a = these are council members who are employed by the municipality they live in.*

The results for the Poirier's partial observability bivariate probit are presented in Table 6. The municipalities are sorted according to their population so that the larger municipality (or

hypothetical municipality) is always in the equation (1) and the smaller municipality is in the equation (2).[9]

Table 6. Results from Poirier's bivariate probit model.

	Smaller		Larger	
	Coeff.	Std. Error	Coeff.	Std. Error
constant	-4.44	6.33	10.69	10.02
economies of scale	-3.49*	1.42	5.61*	2.53
diseconomies of scale	-1.20**	0.37	0.002	0.110
mean of taxable income	-13.09*	5.19	8.93**	3.28
dependency ratio	-38.33	23.11	31.69*	14.76
cooperation	-1.06	0.69	9.89	14.93
merger subsidy	5.99*	2.63	-14.14**	3.64
number of municipalilties	1.03	0.55	-0.26	0.28
distance	-12.87	7.39	-4.66*	2.05
total land area	-58.51**	16.53	22.25**	6.25
deficit	-2.02**	0.66	0.40**	0.35
difference in population	-1.66	0.95	-0.50**	0.12
difference in taxable income	-3.63	2.54	-0.74	2.37
difference in tax rates	42.97	64.77	51.02	44.61
reduction in council size	5.68	3.85	-15.00	4.09
Herfindahl index	10.17*	4.53	16.44*	8.33
same largest party	1.91	1.29	-3.14**	0.79
left-wing party share	13.51**	4.80	-17.00**	5.06
municipal employees' share in council[a]	1.69	5.10	-1.65	8.28
municipal employees' share of population	86.77*	36.59	-35.14*	17.90
Rho (Std. Error)	-0.9998 (0.0011)			
Number of obs.	420			
Log-L	-52.71			

[a] *these are council members who are employed by the municipality they live in. ** and * indicate statistical significance at 1and 5 percent level, respectively.*

Similar to Brasington (1999, 2003a and 2003b), we find a corner solution in a sense that the error term correlation is close to minus unity. Brasington disregards this serious problem. We conduct sensitivity analysis with respect to the correlation coefficient. We impose a grid on the correlation and conduct the estimation within this grid. All the results are quite robust to different values of the correlation coefficient. We suggest using this sensitivity analysis even in models where parameter ρ gains an interior value, because we find using extensive bootstrap analysis, that even in models where the original sample ρ is in the interior, in less

[9] We use BFGS algorithm and the STATA biprobit command to conduct the numerical optimization. We will later use the WESML estimator by Manski and Lerman (1977) to take into account the choice based sampling weights. Also weighting based on the treatment of multiple partner mergers will be conducted later. Besides these sensitivity analyses, we will also use global optimization routines to check whether these possibly local results hold globally.

than 1% of the bootstrapped samples it remained in the interior. Poirier's estimation method is highly unstable and therefore requires careful sensitivity analysis. Another drawback of this approach is that the economic significance of estimates is harder to evaluate than in the logit model.

The reported model is the one with the largest set of explanatory variables that we found while being robust to the sensitivity analysis with respect to error correlation. Although this sort of model selection has a flavour of data mining, it is plausible that it is rather a sign of this specification being a fairly good statistical approximation for this particular data generating process. However, at this point these results should be seen as tentative associations, not causal effects.

The results indicate that smaller municipalities seem to have decreasing returns to scale while the larger municipalities have increasing returns to scale. Economies of scale capture both the economies of scale in service production and the utility losses from not being able to aim the services as precisely in larger municipalities. Therefore this surprising result can indicate Tiebout sorting. A possible story is that the people living in small municipalities want more tailored services whereas those in larger municipalities prefer efficiency. Very poor smaller municipalities seem not to be able to stay independent. For larger municipalities, the income effect is the opposite perhaps indicating their ability to bear the costs of merging with the smaller and often poorer neighbours. Larger municipalities' merger decisions are also associated with higher dependency ratios. Unlike in the logit analysis, the association with merging and being in the same health cooperation is no longer significant. As in the logit model, larger distance hinders mergers but municipalities' land area has intuitive heterogeneous effects. More populated municipalities prefer larger land mass while less populated prefer small land mass. Surprisingly, also deficit has heterogeneous effects. Smaller municipalities seem to like deficits. Perhaps this is an indicator of how used to they are in government bail-outs. The only significant heterogeneity variable is the absolute difference in population. Larger municipalities seem want partners that are of the same size with them. This is opposite to Brasinton’s (2003a) finding and Ellingsen’s (1998) theoretical predictions.

The main variables of interest in this study are variables related local politics. According to the political Herfindahl index it is easier to make merger decisions in a politically concentrated council. Larger councils dislike merger with municipalities where the largest

party is the same as in their council. This could imply that the members of the same party in the other council are perceived as threat in the future elections in the merged municipality. The share of left wing council members has a heterogeneous effect. In smaller councils they seem to support merging and in larger oppose. The share of municipal employees in the council is not associated with merging. Since some these political variables have an effect, the political process seems to hinder the creation of optimal coalitions. This would indicate that the central government should perhaps force coalitions even without the consent of the municipality councils, although the information requirements for the central government for this task are large.

The results concerning the effectiveness of the merger subsidy scheme are very surprising and interesting. We allow ourselves to talk about an effect rather than an association when discussing the effects of government subsidy on mergers since it can be thought of as an exogenous variable. Moreover, we control for the population and the number of disappearing municipalities, which are the only factors affecting the level of subsidies per capita. Furthermore, the step function form of the subsidy scheme should identify its effect from those of population. The per capita merger subsidy had a positive effect on the merger decisions of the smaller municipalities but a negative effect on the large. This means that overall the grant policy is probably a waste even though it works as wanted for the smaller partners. So far we have not come up with a good explanation of why larger municipalities would actually dislike more money. It would have been intuitive if they cared less for the subsidy than the smaller or that they did not consider the subsidies when making the decisions, but a negative effect defies reason. This calls for further work on analysing the reliability of these results.

6 Conclusions

In this paper, we study three important questions concerning municipality mergers in Finland. First, we study the effectiveness of a central government merger subsidy scheme that aimed at promoting municipal mergers. Second, we study the association of local politics with merger decisions. Third, we study how municipal population characteristics, characteristics related to costs of providing statutory services and municipal tax bases are associated with merger decisions.

The most robust results from our coalition level logit analysis are that existing cooperation through a joint authority in producing basic health care has a strong positive effect on the likelihood of merging and that distance prohibits merging. Population of the coalition also seems to matter so that larger coalitions in terms of population are more likely but only up to a point. The optimal municipality population size implied by the logit results is around 60,000.

We also find that local political environment has an effect on mergers. This means that the political process seems to hinder the creation of optimal coalitions. This would indicate that the central government should perhaps force coalitions even without the consent of the municipality councils, although the information requirements for the central government for this task are large. However, due to some econometric problems some of our results should be seen as tentative at this point.

References

Alesina, A. and E. Spolaore (1997): On the Number and Size of Nations. *Quarterly Journal of Economics* 112(4), 1027–1056.

Alesina, A., R. Baqir and C. Hoxby (2004): Political Jurisdictions in Heterogeneous Communities. *Journal of Political Economy* 112, 348–396.

Bradford D. F and W. E. Oates (1974): Suburban Exploitation of Central Cities and Governmental Structure. In H. Hochman and G. Peterson (eds.): *Redistribution through Public Choice*. New York: Columbia University Press.

Brasington, D. (1999): Joint Provision of Public Goods: The Consolidation of School Districts. *Journal of Public Economics* 73, 373–393.

Brasington, D. (2003a): Size and School District Consolidation: Do Opposites Attract? *Economica* 70, 673–690.

Brasington, D. (2003b): Snobbery, Racism, or Mutual Distaste: What Promotes and Hinders Cooperation in Local Public-Good Provision? *Review of Economics and Statistics* 85, 874–883.

Breslow, N. E. (1996): Statistics in Epidemiology: The Case-Control Study. *Journal of the American Statistical Association* 91(433), 14–28.

Butler, J. S. (1996): Estimating the Correlation in Censored Probit Models. *Review of Economics and Statistics* 78(2), 356–358.

Dahlberg, M. and E. Mörk (2006): Public Employment and the Double Role of Bureaucrats. *Public Choice* 126, 387–404.

Gordon, N. and B. Knight (2009): A Spatial Merger Estimator with an Application to School District Consolidation. *Journal of Public Economics* 93, 752–765.

Manski, C. F. (1995): *Identification Problems in the Social Sciences*. Cambridge MA: Harvard University Press.

Manski and Lerman (1977): The Estimation of Choice Probabilities from Choice Based Samples *Econometrica* 45(8), 1977–1988.

Miceli, T. J: (1993): The Decision to Regionalize in the Provision of Education: An Application of the Tiebout Model. *Journal of Urban Economics* 33(3), 344–360.

Poirier, D. (1980): Partial Observability in Bivariate Probit Models. *Journal of Econometrics* 12, 209–217.

Sørensen R. J. (2006): Local Government Consolidations: The Impact of Transition Costs. *Public Choice* 127, 75–95.

Tiebout, C. M. (1956): A Pure Theory of Local Expenditures." *Journal of Political Economy* 64(5), 416–424.

Tiller, K. J. and P. M. Jakus (2005): Applying the Miceli Model to Explain Cooperation in Municipal Solid Waste Management. *Agricultural and Resource Economics Review* 34, 217–225.

Weese, E. (2009): Political Mergers as Coalition Formation: Evidence from Japanese Municipal Amalgamations. MIT mimeo.

Wernicke S. (2005): A Faster Algorithm for Detecting Network Motifs. In S. Istrail, P. Pevzner, and M.Waterman (eds.): *Algorithms in Bioinformatics*. Berlin/Heidelberg: Springer.

Wernicke S. and F. Rasche (2006): FANMOD: a tool for fast network motif detection. *Bioinformatics* 22(9), 1152–1153.

Appendix 1. Maps.

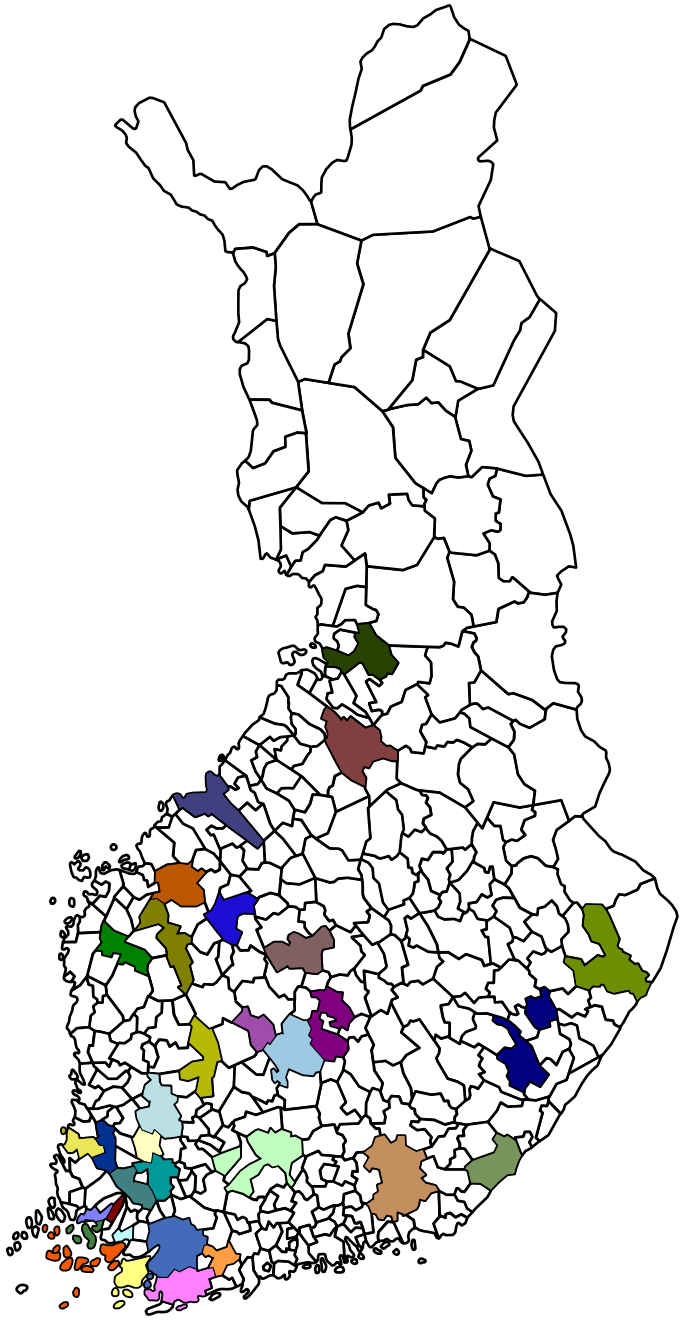

Figure 1A. Map of municipality mergers in 2008–2009.

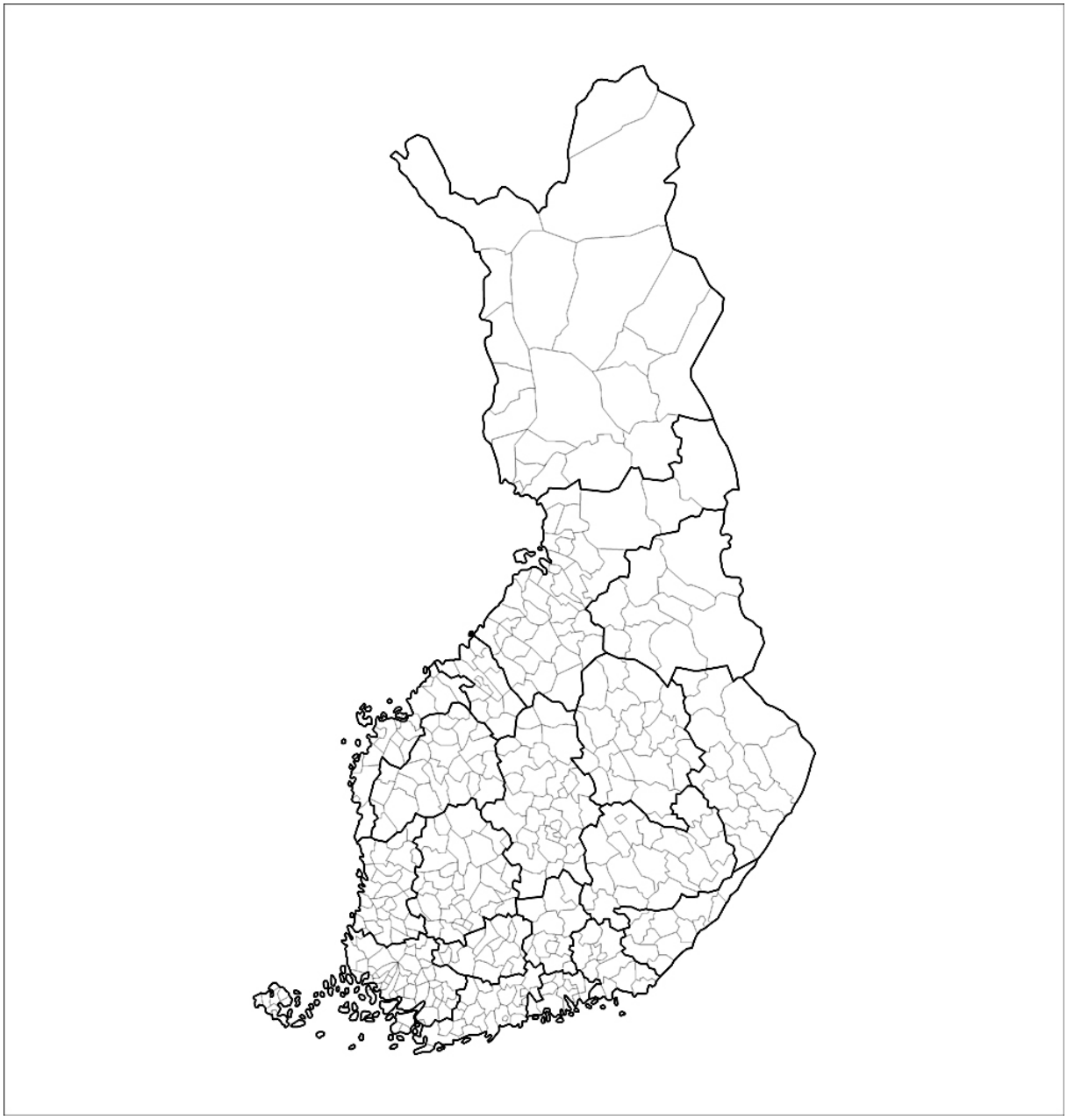

Figure 2A. Map of municipality and county borders, 2007.