SERC DISCUSSION PAPER 102

Politics in Coalition Formation of Local Governments

Tuukka Saarimaa (Government Institute for Economic Research VATT)
Janne Tukiainen (Government Institute for Economic Research VATT)

March 2012

This work is part of the research programme of the independent UK Spatial Economics Research Centre funded by a grant from the Economic and Social Research Council (ESRC), Department for Business, Innovation & Skills (BIS) and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

Politics in Coalition Formation of Local Governments

Tuukka Saarimaa*
Janne Tukiainen*

March 2012

*Government Institute for Economic Research VATT

Acknowledgements
We thank Essi Eerola, Elias Einiö, Ronny Freier, Ari Hyytinen, Aki Kangasharju, Kaisa Kotakorpi, Jani-Petri Laamanen, Teemu Lyytikäinen, Antti Moisio, Henry Overman, Gerard Padro, Rohini Pande, Jaakko Pehkonen, Torsten Persson, Matti Sarvimäki, Otto Toivanen, Eric Weese and an anonymous referee for helpful comments. We also thank seminar and conference participants at EPCS 2010, FEAAM-Tampere 2010, HECER, Jyväskylä, LSE, Manchester, Oxford and VATT for useful discussions. Oskari Harjunen and Niklas Jahnsson provided excellent research assistance. Janne Tukiainen is grateful to Emil Aaltonen Foundation, Jenny and Antti Wihuri Foundation and Yrjö Jahnsson Foundation for financial support and STICERD at the London School of Economics and Political Science for hospitality. Tuukka Saarimaa is grateful to the Finnish Cultural Foundation for financial support and SERC at the London School of Economics and Political Science for hospitality. We also thank Sitra, the Finnish Innovation Fund for funding.

Abstract
We analyze empirically the coalition formation of local governments using a novel reduced form econometric procedure that allows for multi-partner mergers. Using Finnish municipal merger data where mergers were decided independently at the local level, we find that merger decisions are largely in line with voter preferences. Most importantly, mergers are clearly less likely when the distance of the median voter to the coalition centre is large. However, councillors seem also to prefer mergers where post-merger political competition is lower which indicates a concern for re-election. Interestingly, municipalities do not seem to be seeking economies of scale through merging. This is possibly due to existing cooperation in service production which we find to be a strong predictor of merging.

Keywords: Coalition formation, local politics, choice based sampling
JEL Classifications: H77, H72, C35

1 Introduction

The formation and dissolution of nations and local jurisdictions like school districts, municipalities or counties, agreeing on political coalitions, deciding memberships in international cooperative organizations like the UN and WTO and even corporate mergers are all examples of coalition formation. Both history and recent news are full of examples that highlight the practical importance of understanding the mechanisms that, for example, draw the borders between nations. Both the formation and dissolution of Soviet Union is an example of large scale coalition formation and in Belgium the question whether it should be one or two nations is always bubbling under. Due to its significant economic and political importance, coalition formation has attracted interest from both policy makers and theoretical and empirical researchers. The main questions in this literature are how many and what type of coalitions should be optimally formed and how is the optimal distribution of coalitions achieved.

There is a long tradition of economic research concerning local governments. The vast literature on fiscal federalism has concentrated on the question of assignment of functions and taxes to different levels of government.[1] According to Musgrave (1959) and Oates (1972), in the absence of scale benefits in the provision of public goods and interjurisdictional spillovers, decision making should be decentralized so that public good provision can be matched to heterogeneous preferences of the population. A related result due to Tiebout (1956) offers decentralization as a tool for efficient pricing of local public goods in a world where mobile households shop for a suitable tax and public good combination offered by endogenously and optimally formed local jurisdictions.[2] However, these strands of literature leave open the issue of where and through what process local governments or jurisdictions arise in the first place.

Later work on endogenous mergers and coalition formation by e.g. Miceli (1993), Alesina and Spolaore (1997, 2003) and Bolton and Roland (1997) also emphasize the fundamental trade-off between economies of scale in the production of public goods (which favors merging) and regional heterogeneity in preferences (which favors separation).[3] This trade-off determines the optimal size of a given coalition. These studies add to the earlier

[1] See Oates (1999) for an overview of this literature.
[2] See Boadway and Tremblay (2011) for a recent review of the Tiebout model and the subsequent literature.
[3] Dur and Staal (2008) analyze theoretically the role of national transfers in municipal merging.

fiscal federalism literature by introducing a political economy model of coalition formation. Alesina and Spolaore (1997), Bolton and Roland (1997) and Ellingsen (1998) discuss how politics may cause departures from welfare maximizing coalition formation. For example, Alesina and Spolaore (1997) show that the equilibrium under within jurisdiction majority voting is suboptimal to the one achieved under a global social planner.

Empirical contributions to this field are more recent mainly because econometric modeling of spatial merger decisions is complicated for a number of reasons elaborated later on. Brasington (1999, 2003a, 2003b), Gordon and Knight (2009) and Weese (2011) use different methods but are all fundamentally concerned with optimal coalition formation and the characteristics of optimal coalitions in terms of economies scale and preference heterogeneity. Gordon and Knight (2009) and Weese (2011) have proposed structural econometric methods to analyze spatial mergers. All of the existing empirical work abstracts away from a number of interesting political economy questions, such as legislative bargaining and political agency issues.[4] Furthermore, existing structural methods cannot utilize natural experiments for causal inference.

This paper contributes to the literature in a number of ways by analyzing a recent surge of municipality mergers in Finland. First, we introduce a reduced form econometric procedure that allows us to analyze multi-partner mergers and to use standard causal inference tools. Unlike the existing structural methods, this framework also allows for a large set of covariates and statistical inference with a relatively small amount of actual mergers. In a merger analysis, we would like to compare the realized mergers to potential mergers that did not take place. Multi-partner mergers or one-to-many matching causes two related problems for this type of analysis. First, it is more difficult to define the set of potential coalitions that could have formed than it is in a one-to-one matching situation. Second, the number of potential coalitions that can be formed from the original municipal map is very large under most reasonable definitions. We solve these problems using Wernicke's (2005) network detection algorithm which can be used to construct all the potential coalitions that can be formed from the original municipality map while maintaining spatial contiguity.[5] We also illustrate how stratified choice based sampling can easily be incorporated to the analysis in situations where data collection or computational costs prohibit the use of the total population of potential

[4] See e.g. Baron and Ferejohn (1989) and Besley (2006).
[5] Weese (2011) also relies on Wernicke's (2005) algorithm but in a different way.

coalitions, which in many cases can be substantive.[6] This procedure has many attractive features that may be important for future coalition formation studies as well because statistical theory for choice based sampling is well developed and easy to apply. In our application, data collection is not costly with respect to our variables of interest, and thus, we are able to confirm that choice based sampling results are consistent with the results obtained using the full sample where we correct for rare events bias.

Second, in addition to analyzing the usual determinants of coalition formation, such as economies of scale and heterogeneity, we provide the first empirical analysis of the role of local politics in merger decisions.[7] This is possible because municipal councils made all the merger decisions in our data and we observe a number of interesting characteristics of municipalities and their councils. The most interesting aspects of our analysis in this respect are the role of the geographic location of the median voter in a potential coalition, which we can measure quite precisely using a GIS dataset produced by Statistics Finland, and political agency issues related to re-election, political competition and politicians' private careers. The former can be seen as a measure of voter preferences over merging while the latter measures departure from optimality in coalition formation due to politicians' self-interest. Unfortunately, we do not have a natural experiment for most of our explanatory variables. However, descriptive analysis of empirical regularities between political characteristics and merger decisions is novel and therefore interesting as such. Moreover, we can exploit exogenous variation in municipal council size to study the causal effect of political career prospects on merger decisions using regression discontinuity design (RDD). However, the conclusions from the RDD analysis are limited due to simultaneous treatments at the council size thresholds.

Our main result is that measures of voter preferences are the main determinants of coalition formation. Most importantly, the likelihood of merging clearly diminishes as the distance of the median voter to the coalition centre increases. Moreover, more homogenous coalitions in terms of expenditures are more likely to form. However, the local political environment also plays a role. Councilors seem to decide the mergers consistently with

[6] For example, some research questions may require conducting surveys on voter sentiments about various potential coalitions. Such a survey would not be feasible if these voters would have to compare thousands of different potential coalitions.

[7] To our knowledge, only Austin (1999) attempts to look at the political decision making process behind coalition formation. He studies one-sided suburban annexation decisions of US cities in 1950's. However, his variables that measure 'politics' are just measures of population's race, income and tax rate, whereas we observe the actual characteristics of local politicians.

political agency models, potentially indicating concerns over political career. Another interesting observation is that scale economies do not seem to be important. Overall, our results suggest that future theoretical merger analysis should incorporate political agency aspects.

The rest of the paper is organized as follows. Section 2 presents a more detailed description of Finnish municipalities and the institutional background for municipality mergers. In Section 3, we describe our econometric method. Section 4 presents the data and Section 5 the results. Section 6 concludes.

2 Institutional background

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 goods and services are provided by two tiers of government where municipalities constitute the local level.[8] 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. The bulk of Finnish municipalities' expenditures come from producing welfare services with a strong redistributive character, such as social and health care services and primary education. Municipalities fund their spending mostly through their own revenue sources. The most important revenue source is the flat municipal income tax which the municipalities can set freely. However, an extensive central government grant system is used to equalize local cost and revenue disparities. The grant system covers 20 percent of total municipal spending, but for every fourth municipality the grant system covers more than 50 percent of their revenues. Furthermore, cooperation in service production is quite common among municipalities, especially in health care.[9]

Another important feature of the Finnish system is the large number of municipalities relative to population with a large variation in municipal population size. The largest municipality is the country's capital, Helsinki, with almost 600,000 inhabitants, whereas the

[8] See Moisio et al. (2010) for details.

[9] Usually cooperation takes place through a joint authority. Joint authorities are independent legal entities governed by municipal legislation. They are financed by selling their services to municipalities. The most important joint authorities are hospital districts and basic health care. Membership is voluntary, except that every municipality has to belong to a hospital district.

smallest mainland municipality, Suomenniemi, has roughly 800 inhabitants. The median municipal population size is less than 6,000. Despite the size differences, all municipalities have the same responsibilities. Finland is also sparsely populated and population density varies a great deal between municipalities, as do other population characteristics such as age structure and income.

The number of municipalities in Finland has diminished considerable since the 1940's as can be seen from Figure 1. The reduction in the number of municipalities in the 1940's was due to losing land areas to the Soviet Union after the Second World War. There have been two major merger waves. In the 1960's almost all and in the 2000's all of the mergers have been voluntary. Besides the two merger waves, municipality structure has been quite stable through time.

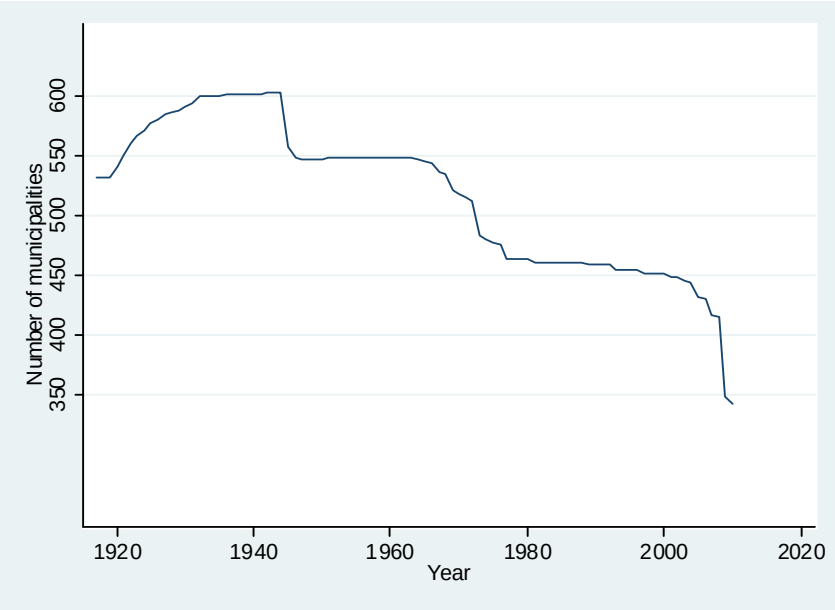

Figure 1. Number of municipalities in Finland, 1917–2010.

What explains the latest drop in the number of municipalities? Some recent developments in Finland have changed the operating environment of municipalities. A major recession in the early 1990s hit different regions with a different force due to differences in industrial composition. The subsequent recovery was regionally uneven and regional disparities started to grow. Furthermore, population aging continues to challenge municipal finance in the future because a bulk of municipal expenditures is related to health and elderly

care. The central government reacted to these concerns in 2005 by initiating a plan that aimed at reforming municipal revenue structure and more importantly making the production of statutory municipal services more efficient. In 2007 these goals were made concrete by a provisional law where the main tool for strengthening the operating environment of municipalities was municipality mergers. The law clearly states that municipalities should have strong enough revenue and labor force bases to cope with the production of statutory public services. According to the law, merging is voluntary and a given merger takes place if it gains the majority vote in all the individual municipality councils contemplating the merger. However, the government initiated a new subsidy scheme to encourage merger activity that applies to mergers taking place from 2008 onwards.[10] In addition to the subsidy, the government guarantees that central government grants are not reduced for five years after the merger even if the new municipality is entitled to a smaller overall grant than the merging municipalities individually. The central government also guaranteed that municipal employees would not lose their jobs for five years after a merger.[11]

Finnish Municipalities are governed by councils that have all the decision making power and decisions are based on simple councilor majority.[12] This is true for merging as well. Municipal councils are elected every four years using open list elections that apply the D'Hondt method. Finland has a multi-party system and currently there are eight parties in the Finnish parliament and these parties also dominate municipal politics. In the 2004 municipal elections the three largest parties (the Centre Party, the Social Democrats and the National Coalition) received roughly 68 percent of the votes with roughly similar shares. In addition to the traditional left- and right-wing division (the Social Democrats and the National Coalition), a somewhat peculiar feature of the Finnish political landscape is the strong support for the Centre Party (formerly known as the Agrarian League). The role of the Centre Party is

[10] The subsidy scheme consists of two parts. The basic part is determined according to the population of the new coalition and the combined population of the participating municipalities except the largest municipality. An additional supplement is paid if the number of municipalities diminishes at least by two and increases as the number of disappearing municipalities increase. The subsidy scheme would potentially allow for RDD to study its effects on merging. Unfortunately, we do not have enough realized mergers sufficiently near any population thresholds of the subsidy system to plausibly identify causal effects.

[11] This has been applied in practice. For example, a typical merger of three municipalities has three persons receiving a city manager's salary for five years.

[12] The council also chooses the municipal executive board that consists of councilors. The composition of the board is based on party shares in the council, i.e. each party in the council get seats in the municipal board according to their share of council seats. Mayors or city managers are not elected officials but civil servants and are of relatively small importance in Finland, but the board plays an important preparatory role in the municipal decision making, including the merger process. See Moisio et al. (2010) for more details on municipal politics.

important in municipality mergers also because they have clearly opposed forced mergers implemented by the central government, whereas the Social Democrats and the National Coalition have been more open to a larger role of the central government in merger decisions.

Each municipality has only one electoral district and no quotas are in place. This goes also for the merged coalitions. Council size is a step function of a municipality's population and is determined by law. The council size varies from 13, for municipalities with population 2,000 or less, up to 85 for municipalities with population 400,000 or more.[13] Thus, in case of a merger of a very large municipality and a very small municipality, it is likely that all the council members in the new coalition will be residents of the former larger municipality. Also the size of the candidate list for each party in elections is a deterministic function of council size since the maximum number of candidates per party is equal to 1.5*(council size). Thus, mergers also reduce the number of candidates and influence political competition in the next elections.

3 Econometric modeling of spatial mergers

3.1 Existing methods

Econometric modeling of spatial merger decisions is complicated for a number of reasons. First, merger decision making is two-sided in the sense that we observe a merger only if all potential merger parties agree to merge. Second, each municipality faces multiple potential merging partners but can merge only once during a given period. It is challenging both to account for all the potential merger partners and to restrict the sum of the probabilities of all the potential mergers to one. 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 consisting of more than two municipalities are possible. This means that the number of potential merger partners and coalitions may be very large and it may be difficult and costly to collect the relevant data.

There are three 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

[13] The council sizes for different populations are the following: population less or equal to 2000 (council size 13, 15 or 17), 2001–4000 (21), 4001–8000 (27), 8001–15,000 (35), 15,001–30,000 (43), 30,001–60,000 (51), 60,001–120,000 (59), 120,001–250,000 (67), 250,001–400,000 (75) and over 400,000 (85).

the two-sided decision making, but not the other features. This approach has been used by Brasington (1999, 2003a and 2003b) in his analysis of school district mergers in the U.S. Second, the matching estimator using simulated method of moments proposed 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. Mergers including multiple municipalities are a prominent feature of our data which makes the Gordon and Knight's estimator infeasible in our case. Gordon and Knight (2009) also apply their estimator to school district mergers in the U.S. Third, Weese (2011) introduces a method of estimating the structural parameters of a political coalition formation game based on simulated maximum likelihood. Weese's estimator allows for all the above features including one-to-many matching, which is accomplished with the application of Wernicke's (2005) network algorithm. Weese (2011) applies his method to municipality mergers in Japan.[14] Due to computational reasons, a number of potentially important factors contributing to merger decisions are left out in both Gordon and Knight's (2009) and Weese's (2011) empirical applications. Moreover, the underlying theoretical models in these studies abstract away from political economy type questions, since they are based on maximizing utility functions of coalitions or individual jurisdictions (or populations of coalitions). For these two reasons, municipal council characteristics cannot be easily included in their method.

3.2 Our procedure

Since one of our main contributions is a rich set of interesting explanatory variables, we will not estimate a computationally cumbersome structural model. Nonetheless, we can account for most of the econometric problems raised earlier. We solve the first problem of two-sided decision making by analyzing coalition level data, which abstracts away from individual municipality choices.[15] This choice imposes an assumption that only coalition level shocks

[14] Alesina et al. (2004) also 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. However, they do not analyze actual merging decisions.

[15] In principle, we could also use the Poirier model, but analyzing multi-partner mergers with a bivariate model would require some unrealistic assumptions. Moreover, the Poirier model typically has bad convergence properties. For example, most of Brasington's (1999, 2003a, 2003b) results are corner solutions to the maximum likelihood optimization problem. According to our tests, convergence is a major problem also in the Poirier analysis of our data.

are relevant or that all the merger parties value the merger in the same way.[16] This may be a potential source of bias in our descriptive analysis, but should not matter for RDD analysis. Unfortunately, we are not able address the second problem of one municipality appearing in many potential coalitions in this paper, and thus, the estimated merger probabilities may be biased. We address the third problem of spatial interdependence by using data only from a single period of the coalition formation game. Assuming that the decision making is simultaneous, the choice set remains stable over our period of analysis. We address the fourth problem using Wernicke's (2005) network detection algorithm. We also introduce a choice based sampling procedure that makes our empirical strategy feasible for a much larger set of questions and applications as elaborated below. We analyze both the full sample using rare events logit and the choice based sample using weighted logit analysis.

The starting point of our procedure is to create a dataset to be analyzed, i.e. the potential coalitions that can be formed from the original municipality map before the mergers. In creating the set of potential coalitions, we follow Weese (2011) and utilize the network algorithm by Wernicke (2005), but in a slightly different way. Using the original municipal map, this algorithm can be used to create all the potential coalitions that could be formed while maintaining spatial contiguity. That is, algorithm takes into account geographic 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 neighbors. This happens when they are a part of a coalition including multiple municipalities that jointly form a geographically coherent municipality.[17] The resulting set of potential coalitions can be either used in our analysis directly or we can sample control coalitions from it.

One practical problem in creating all potential coalitions in our application of municipality data is that large coalitions may take many different geographic shapes and some of the coalitions may be quite unrealistic. 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 borders.[18] The map in Figure

[16] Gordon and Knight (2009) make this assumption throughout their analysis while Weese (2011) is able to relax this assumption. Weese (2011) also shows that relaxing this assumption is important and changes his qualitative results.

[17] We use the FANMOD software introduced in Wernicke and Rasche (2006) to implement the algorithm.

[18] Weese (2011) faces a similar problem because the largest merger that took place during his analysis period in Japan included fifteen municipalities. This creates problems for Weese's estimator because the number of potential coalitions with a size of fifteen or less is huge. Weese circumvents the problem in almost the same way as we do. He limits the analysis to coalitions that cross at the maximum two county borders.

A1 in the Appendix illustrates municipality and county borders. From the map, it's clear that coalitions that stay within county borders (bold line) are much less likely to be problematic in this sense. This is a sensible restriction because none of the actual mergers crossed a county border. 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. Second, county division is almost identical to hospital district division and every municipality has to be part of a hospital district.[19] This makes counties a natural cooperation environment for municipalities even in the absence of formal mergers. Furthermore, because county border would be a perfect predictor of not merging, it would not be possible to use all the coalitions and simply include county border as a right hand side variable. Table 1 presents the number of potential coalitions of different size using the county restriction and the number of actual mergers.

Table 1. Data frequencies and sampling weights by coalition size.

Coalition size	Potential coalitions	Actual mergers	Control coalitions	WESML weights merger = 1	WESML weights merger = 0
2	785	17	340	0.45478	1.02726
3	2,175	8	160	0.07724	1.04614
4	6,561	4	80	0.01280	1.04936
5	20,674	1	20	0.00102	1.04995
6	66,606	2	40	0.00063	1.04997
All	96,801	32	640		

Note: The table presents the total number of potential coalitions, the number of mergers that actually took place, the sampled control coalitions and the WESML weights.

Although not an issue in our application, in a large map with mergers involving many municipalities, the set of potential mergers would be too large for any computational analysis. Using choice based sampling the data set can be diminished to a manageable size while maintaining valid statistical inference. Moreover, some research questions may involve costly data collection and sampling could be done prior to collecting more data to limit the costs.

[19] We also extract the counties of Kainuu and Lapland from our analysis. Kainuu County is currently experimenting with a county level council, and thus, the municipalities will not merge while the experiment is ongoing. Lapland is an outlier in the data, with large land area and low population municipalities. Therefore, drawing control coalitions from Lapland would result in inefficient sampling. Moreover, there were no mergers in Lapland during our analysis period. The map in the Appendix highlights these counties.

Using our data, we can verify that analyzing a choice based sample will replicate the full sample results.

In choice based sampling, the basic idea is to keep all the rare events in the data as a treatment group. In our case, we include all the actual mergers, and then randomly sample a control group from the common events, which are the potential coalitions that did not merge.[20] One way to interpret this set up is that the population of interest is the set consisting of all potential coalitions. Choice based sampling can be made more efficient by conditioning it on observed characteristics of the treatment group. Sampling conditioned on explanatory variables is typically referred to as exogenous stratified sampling while choice based sampling is an example of endogenous sampling. Combination of these two is often referred to as stratified choice based sampling. This sampling method has been widely applied especially in epidemiology with rare diseases, but it is also common in econometrics.[21] Our sampling procedure is an extension to the existing literature on empirical spatial merger analysis. It reduces the computational burden dramatically and can be used in both reduced form and structural analysis because the statistical theory on choice based sampling is well developed and estimators that handle data from this type of sampling are readily available.[22]

The sampling procedure is as follows. First, we select all the actual mergers into our sample. This type of endogenous sampling is necessary because a random sample from all potential coalitions would consist only of zeros with a high probability as can be seen from Table 1. For each actual merger in the data of size n (n = number of merging municipalities), we sample 20 potential coalitions of equal size that did not actually take place as controls.[23] We repeat this for all n that actually took place (in our case 2, 3, 4, 5 and 6).[24] Second, we stratify the sampling on the number of merging municipalities. Otherwise, the control group would consist almost only of mergers with many partners making the estimation very inefficient. This type of stratification is standard practice in choice based sampling.

[20] We use the term control group to follow the terminology of case-control sampling. This does not imply any causal analysis as such.
[21] See e.g. Breslow (1996) and Manski (1995).
[22] See e.g. Manski and Lerman (1977), Imbens (1992) and Breslow (1996).
[23] The number of controls should be decreasing in the costs of collecting the variables of interest. In the logit model, as large a number as possible is preferred, since rare events correction is possible (see King and Zeng 2001) but in some other models, a balanced smaller sample may be preferred over a larger non-balanced sample due to convergence issues.
[24] We omit a single merger involving 10 municipalities as an outlier and include a dummy for the coalitions in the county of the omitted coalition.

Stratified choice based sampling is prone to two sources of bias that we need to account for.[25] First, the analysis needs to be conditioned on strata fixed effects. In our case, the strata groups are defined by the different number of merger partners (coalition size). Failure to do so results in omitted variable bias.[26] Second, we need to use weights based on the sampling rates, which can be done by using the weighted exogenous sampling maximum likelihood estimator (WESML) proposed by Manski and Lerman (1977). WESML weights can be applied with any maximum likelihood estimation method and in our case the weights can be calculated using the numbers in Table 1.[27] Within each stratum, the weights take two values (w_1/p_1) and (w_0/p_0), where w_1 and w_0 are the population proportions of merging and not merging and p_1 and p_0 are the respective sample proportions.

One of the main advantages of our procedure is that usual methods of causal inference, such as RDD, instrumental variables and difference-in-difference methods, can be used without much further complications. This is obvious for the full sample but less clear for the choice based sample. As an example, consider the case of RDD. Obviously, if stratification is not correlated with the selection into the treatment, regression discontinuity treatment effect is consistent. However, if stratification is correlated with the selection into the treatment the issue is not as clear. Let us consider the worst case scenario of perfect correlation with an abstract example. Assume that we stratify the sampling based on coalition population being under or above 20,000 and the discontinuity of interest is also at 20,000. Due to a balanced stratification scheme, this would result in having exactly the same relation of ones and zeros in the outcome variable in either side of the discontinuity threshold and on average this identical relation would persist near the discontinuity. If we estimated the model without weighting, the treatment effect would always be zero. Similar scenarios can be contemplated where this sort of correlation would lead to either upward or downward biased treatment effects. Fortunately, by using the WESML weights, the population level group effects are estimated consistently. Thus, even with perfect correlation between the selection into the treatment and the stratification, weighting allows for consistent estimation of the treatment effect in RDD.

[25] See e.g. Cram et al. (2010).
[26] For example, transaction and/or negotiation costs are probably increasing in coalition size, which in turn is correlated with a number of observable variables of interest. In fact, in a nonlinear model stratification needs to be controlled for even when stratification is not correlated with observables.
[27] In STATA, this can be done either using the pweight option or the svy survey package. Imbens (1992) provides a more general semiparametric estimator based on the method of moments estimation.

4. Data description

We have linked data on municipal mergers from 2008–2009 to municipal characteristics obtained from Statistics Finland and to characteristics of municipal councilors obtained from the Ministry of Justice and the Local Government Pensions Institution. The councils in our data were elected in 2004 for a four year term and they made all the decisions regarding the mergers that took place in 2008–2009. Future mergers are decided by councils elected in 2008 and after. The new municipal division to take place after the 2008 and 2009 mergers were implemented was used already in the October 2008 elections. The 2004 councils decided also 2007 mergers but they were decided under a very different merger subsidy scheme than the 2008–2009 mergers. By 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 where municipalities make simultaneous choices over the same map.

Some descriptive statistics of our coalition level data are presented in Table 2, where we compare the actual mergers to all potential coalitions that could have merged and to our stratified choice based sample from this population. All council characteristics are calculated using 2004 election data and all coalition characteristics are based on 2007 municipality data. Variable definitions are presented in Table A1 in Appendix A. Actual mergers differ from the potential coalitions in a number of respects whereas the choice based sample is much more similar probably due to stratification. Especially striking is that actual mergers are clearly more likely to have cooperated already before the merger. Existing cooperation here means that the municipalities in the coalition had a joint authority in producing basic health care services, which is the largest single expenditure item for municipalities.[28]

In actual mergers, the distance of the median voter to the centre of the largest municipality in the coalition is much lower than in all potential coalitions or the choice based sample. The median voter distance measure is calculated using Geographic Information System (GIS) techniques in the following way. We have data on the location of eligible voter population within 250 m * 250 m grids for the whole of Finland, which we use to calculate a Euclidean distance from all the grids within a coalition to the centre of the largest municipality in the coalition. Since we have data on the number of eligible voters in the grids

[28] Cooperation was in place long before these mergers were contemplated. Thus, cooperation can be treated as an explanatory variable rather than an alternative response variable.

we can calculate the median distance. The population of the largest municipality is excluded from these calculations. The intuition is that if a merger goes through it is likely that some, if not all, municipal services will be concentrated into the largest municipality, and thus, the distance to services will grow for voters living in smaller municipalities in a given coalition. This is a sensible measure because any single municipality in a given coalition can prevent a merger.

Actual mergers are also much less heterogeneous when compared to potential coalitions. In part this is due to the fact that larger coalitions are more heterogeneous simply because there are more municipalities involved and that large coalitions dominate our data. For example, the expected maximum difference in municipal characteristics is larger for coalition of size six than for a coalition of size three. This is obvious when one looks at the heterogeneity measures for the choice based sample which are much closer to the actual mergers.

The importance of local politics is captured with variables that measure political fragmentation (the Herfindahl index), similarity of political power in different municipalities and party composition of the councils. Moreover, we aim to capture the importance of politicians' private incentives in merger decisions with variables that measure the importance of both political and private career prospects. We measure political career prospects using three variables that are all based on municipal council size set by law. Our first variable is the "Reduction in council size", which measures the relative amount of council seats lost if the merger goes through. The larger this variable is the more seats are lost in relative terms and it should capture changes in councilors' re-election probabilities and also the relative power that they will have if they are re-elected into to new council. Thus, the variable measures the expected benefits of being a councilor if a merger goes through. However, we cannot identify re-election and political power considerations separately.

In addition, we measure changes in political competition using two variables based on party list size. According to law, the maximum number of candidates a party is allowed to nominate is equal to 1.5 times the council size. Using this rule, we calculated for each coalition the share of both 2004 candidates and councilors that would not fit into their party's candidate list if a merger goes through. That is, we measure the share of candidates and councilors that would not be able to run in the next elections simply because of the list size rule. These variables are denoted as "Party list size candidates" and "Party list size

councilors", respectively. The rationale for using two variables is that councilors who decide the mergers may be concerned with competition that comes from insiders, that is, the current councilors who decide the mergers, or from outsiders, which is the pool of all candidates most of whom cannot vote for a merger because they are not sitting in the current council. As can be seen from Table 2, all these variables are clearly smaller for the actual mergers compared to all potential coalitions, but the differences are small between mergers and the choice based sample.

Finally, we are interested whether councilors' private career prospects matter for merger decisions. Being a member of a municipal council in Finland is a part-time job that incurs practically no salary even in the largest municipalities. However, some council members are municipal employees (as teachers, doctors, bureaucrats etc.), which may affect their attitude toward mergers, since cost savings in public service production are often put forth as the primary reason for mergers. The central government anticipated some resistance from municipal employees and enacted a transition period spanning five years after a merger during which municipal employees cannot be laid off due to merger-related efficiency reasons. Councilors' private career prospects are captured with the variable "Municipal employees share in council". This variable may, of course, also capture voter sentiments because a high share of municipal employees in the population could be correlated with the council share. In order to identify the effect of the potential merger on council member's own career separately from its effect on their voters' careers, we also include the share of municipal employees in the population to the model.

Table 2. Descriptive statistics.

	Merger = 1		All coalitions Merger = 0		Choice based sample Merger = 0	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Number of observations	32		96,769		640	
Average coalition size	2.84	1.17	5.55	0.78	2.84	1.15
Total population	33,882	33,546	93,695	103,554	42,191	63,647
Mean of taxable income (€)	12,208	1,594	12,164	1,612	11,719	1,742
Total expenditures (€ per capita)	5,475	521	5,604	520	5,504	557
Merger subsidy (€ per capita)	328	216	250	176	358	312
Cooperation	0.50	0.51	0.01	0.09	0.16	0.37
Median voter distance (km)	19.9	14.8	34.9	18.3	28.6	15.8
Language	0.09	0.30	0.17	0.37	0.10	0.30
Population size heterogeneity	22,329	26,789	57,994	75,623	25,449	46,751
Income heterogeneity	2,261	1,424	3,438	1,509	1,997	1,450
Tax rate heterogeneity	0.85	0.62	1.28	0.60	0.73	0.56
Expenditure heterogeneity	819	715	1,717	734	981	735
Largest municipality vote share	0.72	0.17	0.51	0.20	0.65	0.17
Herfindahl index	0.31	0.13	0.30	0.09	0.32	0.10
Same largest party	0.66	0.48	0.46	0.50	0.65	0.48
Centre Party share	0.34	0.20	0.40	0.17	0.40	0.18
Left-wing party share	0.31	0.12	0.30	0.09	0.29	0.10
National Coalition Party share	0.19	0.09	0.19	0.08	0.18	0.09
Other party share	0.16	0.18	0.11	0.08	0.13	0.14
Reduction in council size	0.44	0.12	0.67	0.06	0.45	0.13
Party list size candidates	0.15	0.13	0.33	0.08	0.14	0.11
Party list size councilors	0.01	0.06	0.04	0.07	0.01	0.03
Mun. employees share council	0.22	0.06	0.21	0.03	0.22	0.05
Mun. employees share population	0.06	0.01	0.06	0.01	0.06	0.01

5 Econometric results

5.1 Descriptive regression results

In this section, we present two sets of econometric results. The starting point for our first approach is the theoretical argument that municipalities seek economies of scale through merging but face a trade-off in terms of matching service production to more heterogeneous preferences of a larger population. Furthermore, we assess the role of local politics in coalition formation. In the second approach, we test for political agency issues using discontinuities in council size as our natural experiment.

The estimation procedure is straightforward. Using the binary logit model we study which types of coalitions are likely to form. This approach can be seen as estimating a

coalition formation game in reduced form. Let y_k denote the merger decision of coalition *k,* i.e. y_k equals one if the coalition forms an actual new municipality and zero otherwise. The logit model can be written as

$$(1) \qquad y_k = 1\left(\mathbf{x}_{1k}'\boldsymbol{\beta}_1 + \mathbf{x}_{2k}'\boldsymbol{\beta}_2 + \mathbf{z}_k'\boldsymbol{\delta} + f_k + \varepsilon_k > 0\right),$$

where 1(.) is an indicator function that equals one if the statement in the parentheses is true and zero otherwise. The model includes three sets of covariates along with stratification group fixed effects, f_k. The vector $\mathbf{x}_1$ includes coalition characteristics (and a constant), $\mathbf{x}_2$ is a vector of variables describing coalition heterogeneity, $\mathbf{z}$ a vector of municipal council characteristics at the coalition level and ε_k is the usual logit error term.

Our main results are presented in Table 3. The first model in Table 3 is a rare events logit model estimated using the whole data.[29] The second model is a standard logit model using the choice based sampled data and WESML weights. These two methods are both consistent and should produce fairly similar results. For comparison, we also report results from a linear probability model estimated using OLS and the whole data. Table 4 present robustness results using different subsets of variables from the richest model specification in Table 3. All the continuous variables in the models are standardized to have a zero mean and a standard deviation of one. Although we have a rich set of covariates, there is always the possibility that we have overlooked or do not observe some important variables that drive merging. Since it is unlikely that all the unobservable factors driving the mergers are independent of the observable ones, the results should be interpreted cautiously as associations rather than causal effects. Possible exceptions are the merger subsidy, the reduction in council size and party list measures because they are exogenous step functions based on municipal population and we control for population and its square term directly and include coalition size fixed effects. However, we remain cautious in interpreting the results concerning these variables as causal.

[29] We use the Relogit command in STATA. See King and Zeng (2001) and King and Zeng (2003) for details. Not correcting for the rare events would result in underestimation of the merger probability.

Table 3. Results for rare events logit, WESML logit and OLS.

	Rare events logit		Logit, WESML		OLS	
	Coeff.	**Std. Error**	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**
Constant	-8.724***	2.616	-9.556***	2.628	0.016***	0.005
Population	-1.802	5.422	-8.569	6.009	-0.001	0.0007
$(\text{Population})^2$	0.413	1.115	-2.542	2.740	0.0001***	0.00004
Mean of taxable income	0.698	0.495	0.681	0.544	0.0002	0.0003
Total expenditures	0.293	0.254	0.524	0.322	-0.00003	0.0002
Merger subsidy	0.237**	0.103	0.067	0.160	-0.0002	0.0002
Cooperation	1.313**	0.580	0.663	0.731	0.013***	0.0043
Median voter distance	-1.188*	0.638	-1.579**	0.675	-0.0002*	0.0002
$(\text{Median voter distance})^2$	0.270***	0.067	0.539***	0.186	0.0001**	0.00005
Language	-0.371	0.869	-0.052	0.989	-0.0004	0.0006
Population size heterogeneity	-0.711	4.798	4.054	4.938	-0.0004	0.0006
Income heterogeneity	-0.085	0.342	0.306	0.562	0.00005	0.0001
Tax rate heterogeneity	0.265	0.288	1.022***	0.365	0.00004	0.0001
Expenditure heterogeneity	-0.831**	0.354	-1.315***	0.460	-0.0002*	0.0001
Largest municipality vote share	1.707**	0.761	1.274	0.784	0.0006*	0.0003
Herfindahl index	-0.358	0.346	-0.330	0.517	0.0003	0.0007
Same largest party	0.875	0.556	2.601***	0.895	0.0008*	0.0005
Centre Party share	-0.012	0.546	-0.219	0.707	-0.0007	0.0006
Left-wing party share	0.109	0.336	0.333	0.474	0.0002	0.0003
Other party share	0.132	0.245	-0.423	0.311	0.0002	0.0004
Reduction in council size	-0.101	0.210	0.169	0.232	-0.00003	0.0004
Party list size candidates	0.573*	0.295	-0.223	0.469	0.0002	0.0002
Party list size councilors	1.127*	0.645	3.187***	0.727	0.0005	0.0003
Mun. employees share council	-0.047	0.145	-0.086	0.166	-0.00002	0.0001
Mun. employees share population	0.069	0.280	0.573	0.467	0.00003	0.0002
Number of obs.	96,801		672		96,801	
Pseudo Log-L			-30.3			
Pseudo R^2 / R^2			0.33		0.02	

Notes: The results are from coalition level models where the dependent variable equals one if the coalition actually underwent a merger. Coalition size fixed effects are included in all models. OLS standard errors are robust to heteroscedasticity. ***, ** and * indicate statistical significance at 1, 5 and 10 percent level, respectively.

The results are interesting in a number of ways and most importantly, the logit model with choice based sample data and WESML weights produces results that are consistent with the rare events logit results. In these two models, none of the variables have opposite signs with statistically significant coefficients, although some variables are significant in the rare events logit and not in the CBS logit and *vice versa*. Also the OLS results are largely in line with both of the logit models.

The results indicate that coalition population does not seem to matter, which would indicate that municipalities are not looking for economies of scale through merging. Median voter distance, on the other hand, is an important driver of merging, although the negative effect diminishes as distance grows.[30] Existing cooperation is also important and increases the likelihood of merging. This a very interesting result because it indicates that merging takes place between municipalities that may already benefit from economies of scale in producing basic health care services. This casts some doubts on whether these mergers lead to increased cost-efficiency in service production. Furthermore, voters in cooperating municipalities are consuming health care services of the same quality, and thus, probably have similar preferences for these services. Expenditure heterogeneity is negatively associated with merging, which means that municipalities with heterogeneous preferences for municipal spending are less likely to merge. These results clearly indicate that preference heterogeneity is an important factor behind merging and support the predictions by Alesina and Spolaore (1997). Somewhat surprisingly though, coalition level mean income and within coalition income heterogeneity are not statistically significant. The central government merger subsidy gets the expected positive sign but it is statistically significant only in the rare events logit.

Next we turn to results concerning politics. At coalition level, political parties do not seem to matter. All the party share variables, the Herfindahl index and the same largest party are not statistically significant in the rare events logit, although the same largest party is positive and significant in the WESML logit. These results are not very surprising because in Table 3 we are controlling for heterogeneity in expenditures and tax rates directly and it can be argued that variables based party shares also reflect voter preferences. Furthermore, these results remain unchanged even without heterogeneity measures as can be seen from Table 4. Interestingly, the vote share of the largest municipality gets a positive sign and is statistically significant. It may be that councilors of larger municipalities who are elected in more competitive elections, value future political careers more than councilors of smaller municipalities while these smaller municipalities are more often forced to merge for economic reasons.

The results so far suggest that municipal councilors' actions are in line with voter preferences. What about the variables measuring political and private career concerns? Reduction in council size is not statistically significant. However, the party list size variables

[30] For a large majority of the data, the negative effect clearly dominates.

that measure political competition get a positive sign and are statistically significant at 10 percent level. In fact, the *p*-value for the Party list size councilors is 0.052 and they are jointly significant at 5 percent level. Once we control for the reduction in council size these variables have a clear interpretation as measures of political competition related to re-election. The result means that coalitions with lower political competition are more likely to form. The share of municipal employees in council does not seem to matter for merging.

Table 4. Additional rare events logit results.

	Rare events logit		**Rare events logit**		**Rare events logit**		**Rare events logit**	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
Constant	-5.144***	1.165	-7.194***	1.736	-7.341***	1.940	-5.045***	1.196
Population	-0.566	0.983	-7.729	4.873	-2.367	1.275		
$(\text{Population})^2$	-0.457	0.915	-1.335	0.978	0.200	0.985		
Mean of taxable income	0.422	0.384	0.643	0.464	0.445	0.386		
Total expenditures	0.218	0.165	0.426**	0.190	-0.011	0.241		
Merger subsidy	0.037	0.170	0.061	0.158	0.232*	0.123		
Cooperation	1.538***	0.513	1.351***	0.490	1.356**	0.558		
Median voter distance	-0.775	0.510	-1.028**	0.512	-1.013*	0.593		
$(\text{Median voter distance})^2$	0.253***	0.061	0.292***	0.058	0.264***	0.066		
Language			-0.215	0.677				
Population size heterogeneity			5.978*	3.497				
Income heterogeneity			-0.097	0.338				
Tax rate heterogeneity			0.165	0.253				
Expenditure heterogeneity			-0.796**	0.348				
Largest municipality vote share					1.543***	0.435	0.829***	0.280
Herfindahl index					-0.300	0.314	0.361	0.283
Same largest party					0.790	0.559	0.607	0.521
Centre Party share					0.076	0.570	-0.361	0.555
Left-wing party share					0.195	0.364	0.208	0.404
Other party share					0.163	0.247	-0.037	0.276
Reduction in council size					-0.088	0.195	-0.053	0.187
Party list size candidates					0.604**	0.289	0.103	0.306
Party list size councilors					0.961	0.621	0.746	0.534
Mun. employees share council					-0.012	0.149	0.071	0.145
Mun. employees share population					0.017	0.252	-0.401**	0.171
Number of obs.	96,801		96,801		96,801		96,801	

Notes: The results are from coalition level models where the dependent variable equals one if the coalition actually underwent a merger. Coalition size fixed effects are included in all models. ***, ** and * indicate statistical significance at 1, 5 and 10 percent level, respectively.

5.2 Council size RDD results

As we argued above, the reduction in council size and party list measures are exogenous step functions based on municipal population and as such the results may be interpreted as causal since we are controlling for population with second order polynomial. However, as an additional test for political agency issues we perform an RDD analysis using discontinuities in council size as our natural experiment. The problem in using these discontinuities is that a number of things change at the thresholds and it is impossible to identify which of the channels produces the possible treatment effect. However, even in this case RDD can be used to determine whether the simultaneous treatments have a significant joint effect and whether some of the effects dominate others.

If the total population in a new merged municipality lands just above a population threshold, there are more seats available and re-election is easier. If re-election is important for the councilors, they should favor mergers that happen to land just above a threshold over mergers that land just below. On the other hand, if councilors prefer power per capita (or councilor) over more certain re-election, they may favor mergers just below a threshold. Thus, the treatment affects both the probability of re-election and the benefits of office conditional on being re-elected, and these two effects should have the opposite sign.

Crossing the population threshold also affects the number of candidates that each party is allowed to nominate in the next elections because the size of the party list is limited to 1.5 times the council size. This limit is typically not binding in smaller municipalities but it is binding in larger ones. In particular, in coalitions with a large population or many merger partners, this rule significantly limits the number of candidates that parties can nominate if a merger goes through. As a result, incumbent councilors face more competition in the next elections if the new coalition is above a population threshold, and thus, if incumbents care for re-election, they should prefer coalitions just below the threshold. RDD based on the council size rule only identifies the joint effect of three simultaneous treatments, two of which have opposite effects on re-election probabilities, and one with a positive effect on the benefits of office conditional on re-election.[31] The fact that the council size rule implements simultaneous treatments also means that council size thresholds cannot be used as instrumental variables because we would have more endogenous variables than instruments.

[31] Only council size and maximum party list size change at the threshold. The test for balance of other covariates around the thresholds does not dispute this.

Some concerns have been raised recently, by e.g. Ade and Freier (2011), about the validity of using population thresholds in RDD. In addition to the discussed problem of simultaneous exogenous treatments, they point out two other typical issues and show that all three invalidate some existing published research.[32] First, additional endogenous choices on other institutions are often taken simultaneously to the response of interest. In our case, no other decisions are made simultaneously to the merger. Second, manipulation and precise control over population measures would invalidate the experiment. In our set up, the manipulation of population statistics would be very costly for the municipalities because this information is gathered independently by the Central Government. [33] Moreover, the potential 'manipulation' of resulting true population of the realized mergers is the variable of interest for us.

Due to a small number of realized mergers in our data, we are forced to use parametric RDD and the identification of the treatment effect depends on the correct specification of the forcing variable. We include many polynomials of population in the regression and test for robustness of the results with respect to changing the number of polynomials. Furthermore, we alleviate this identification assumption by comparing means only near a threshold by including both a dummy for a band around the threshold and a treatment dummy indicating that the observation is above and near a threshold. We test for robustness with respect to the size of the band as well. Since the council size rule has multiple discontinuities we average over the treatments at different thresholds by including only one treatment and band dummy in each regression. Furthermore, we use the actual population as the forcing variable instead of distance to the nearest threshold. This is important because we have such a small sample of mergers and we do not want to compare coalitions just below a population threshold of 2,000 to those just above a 250,000 threshold for example.

In addition to a merger RDD we also report regressions that illustrate the effect of crossing a threshold on our political career variables used in the descriptive regressions. Thus, we run regressions of the type:

$$(2) \qquad y_k = \beta_0 + \beta_1 band_k + \beta_2 treat_k + g\left(pop_k\right) + f_k + \varepsilon_k,$$

[32] E.g. Egger and Koethenbuerger (2010).
[33] We also conducted tests for manipulation of the forcing variable. According to the McCrary test, our experiment seems to be valid.

where y_k is the outcome in question (merging or one of the political career variables), $band_k$ is an indicator for whether coalition k is within a bandwidth of any threshold and $treat_k$ is an indicator for whether coalition k is within a bandwidth and above any threshold. The function $g(pop_k)$ is a high order polynomial of coalition population and f_k denotes coalition size fixed effect.

The results are presented in Table 5. First, it is clear that the council size treatment effect for all the variables related to political agency issues is large and significant. These results confirm that RDD based on the council size rule is a good natural experiment for the joint impact of these variables. The result also validates the argument above that the council size thresholds cannot be used as instrumental variables. However, we do not find any significant effects of crossing the threshold on merger activity. The result is robust to different band sizes and number of polynomials.[34] Thus, we are not able to confirm that political agency issues are relevant for the merger activity. On the other hand, we cannot reject their possible influence either. Naturally, some caution is in order because the estimates are not extremely precise in all cases. Again that the RDD results from the rare events logit model are similar to the WESML logit model using the choice based sampled data.

[34] The results are robust to band sizes of 10 %, 15 %, 20 % and 25 %, and 0–7 polynomials of coalition population. Smaller bandwidths are not feasible due to limited amount of actual mergers around the thresholds. The results are also robust to adding control variables.

Table 5. Council size RDD results.

	All data, N = 96,801			CBS, N = 672		
Forcing variable polynomials	Squared	Cubic	Quartic	Squared	Cubic	Quartic
			Reduction in council size			
Window size 15 %	-0.637***	-0.648***	-0.652***	-1.645***	-1.645***	-1.669***
	(0.006)	(0.006)	(0.006)	(0.141)	(0.141)	(0.141)
Window size 20 %	-0.522***	-0.539***	-0.545***	-1.379***	-1.377***	-1.403***
	(0.005)	(0.005)	(0.005)	(0.132)	(0.134)	(0.133)
			Partry list candidates			
Window size 15 %	-0.596***	-0.593***	-0.583***	-0.758***	-0.746***	-0.755***
	(0.008)	(0.008)	(0.008)	(0.101)	(0.101)	(0.102)
Window size 20 %	-0.499***	-0.497***	-0.487***	-0.576***	-0.565***	-0.575***
	(0.007)	(0.007)	(0.007)	(0.091)	(0.092)	(0.092)
			Partry list councilors			
Window size 15 %	-0.280***	-0.307***	-0.344***	-0.103**	-0.106**	-0.108**
	(0.008)	(0.008)	(0.008)	(0.048)	(0.049)	(0.049)
Window size 20 %	-0.284***	-0.322***	-0.364***	-0.055	-0.058	-0.059
	(0.007)	(0.007)	(0.007)	(0.039)	(0.040)	(0.039)
			Merger			
Window size 15 %	0.060	0.048	0.027	0.029	-0.339	-0.377
	(0.528)	(0.529)	(0.525)	(0.525)	(0.730)	(0.678)
Window size 20 %	0.155	0.138	0.116	0.117	-0.176	-0.204
	(0.466)	(0.468)	(0.470)	(0.470)	(0.628)	(0.596)

Notes: The regressions for council size and party list variables are estimated using OLS. Merger regressions are estimated using rare events logit and WESML logit. Standard errors are reported in the parentheses. In the OLS regressions, the standard errors are robust to heteroscedasticity. All regressions include coalition size fixed effects. ***, ** and * indicate statistical significance at 1, 5 and 10 percent level, respectively.

6 Conclusions

This paper analyzed empirically the coalition formation of local governments. We introduced a novel econometric strategy involving choice based sampling from a spatial network to handle multi-partner mergers and applied the method to recent municipality mergers in Finland. We analyzed coalition formation using logit analysis where a coalition consisting of two or more municipalities is used as the observation unit. Our main interest was the association of local politics with merger decisions, a clear gap in both the empirical and theoretical literature on local government coalition formation. The most interesting aspects of our analysis in this respect are the role of the geographic location of the median voter in a potential coalition, which we can measure quite precisely using a GIS dataset produced by

Statistics Finland, and political agency issues related to re-election, political competition and politicians' private careers. The former can be seen as a measure of voter preferences over merging while the latter measures departure from optimality in coalition formation due to politicians' self-interest. If political agency considerations matter, optimal coalition formation may be hindered further from the majority voting result presented by Alesina and Spolaore (1997). We also studied how different municipality characteristics and heterogeneity in these characteristics within a potential coalition are associated with merging.

The main result from our analyses is that the councilors' decisions are largely in line with voter preferences. Mergers are clearly less likely when the distance of the median voter to the coalition centre is large. Also within coalition per capita expenditure heterogeneity was negatively associated with merging while existing cooperation in producing basic health care services clearly increased merging. Interestingly, coalition population was irrelevant to merging indicating that municipalities do not seem to be seeking economies of scale through merging. However, the local political environment has a role in merger decisions. Councilors seem to be concerned over re-election and prefer mergers where post-merger political competition is lower. This result is based on an exogenous council size and party list size rule based on municipal population and it is difficult to come up with a credible story where this result would reflect voter sentiments in any way. This means that local political environment may hinder the formation of optimal coalitions. Our results imply that theoretical and empirical work on coalition formation that abstracts away from political decision making process and decision makers' motives may be too simplistic.

The results also raise the question of which government level should make the decisions concerning mergers. One option is that the central government should force mergers even without the approval of the municipality councils. However, there are some obvious caveats to this. First, information requirements for the central government to form optimal coalitions are large. Second, it may not be politically feasible. Third, it is not clear that the central government decision makers do not have private incentives at stake.

Besides providing important information on coalition formation mechanisms in general, our application of municipal mergers should be of practical interest to countries with a multilayer system of government. There have been a number of recent examples of municipality mergers, such as in Germany and Japan in the late 1990's and 2000's and in Denmark in 2007. In all these cases, the merger decisions were mostly made at the local level.

Furthermore, the results 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. For example, the finding that existing cooperation in basic health care is an important factor driving merger decisions could mean that a merger does not produce large efficiency gains because economies of scale are already exhausted through cooperation. However, we leave these issues for further research.

References

Ade, F. – Freier, R. (2011). When Can We Trust Population Thresholds in Regression Discontinuity Designs? DIW Berlin Discussion Paper No. 1136.

Austin, A. (1999): Politics vs. Economics: Evidence from Municipal Annexation. *Journal of Urban Economics* 45(3), 501-532.

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

Alesina, A. – Spolaore, E. (2003): *The Size of Nations*. Cambridge: MIT Press.

Alesina, A. – Baqir, R. – Hoxby, C. (2004): Political Jurisdictions in Heterogeneous Communities. *Journal of Political Economy* 112(2), 348–396.

Baron, D.P. – Ferejohn. J. (1989): Bargaining in Legislatures. *American Political Science Review* 83 (4), 1181–1206.

Besley, T. (2006): *Principled Agents? The Political Economy of Good Government*. Oxford: Oxford University Press.

Bolton, P. – Roland, R. (1997): The Breakup of Nations: A Political Economy Analysis. *Quarterly Journal of Economics* 112(4), 1057–1090.

Brasington, D. (1999): Joint Provision of Public Goods: The Consolidation of School Districts. *Journal of Public Economics* 73(3), 373–393.

Brasington, D. (2003a): Size and School District Consolidation: Do Opposites Attract? *Economica* 70(280), 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(4), 874–883.

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

Cram, D. – Karan, V. – Stuart, I. (2010): Three Threats to Validity of Choice-based and Matched-Sample Studies in Accounting Research. *Contemporary Accounting Research* 26, 477–516.

Dur, R. – Staal, K. (2008): Local Public Good Provision, Municipal Consolidation, and National Transfers. *Regional Science and Urban Economics* 38(2), pp. 160–173.

Egger, P. – Koethenbuerger, M. (2010): Government Spending and Legislative Organization: Quasi-Experimental Evidence from Germany. *American Economic Journal: Applied Economics* 2(4), 200–212.

Ellingsen, T. (1998): Externalities vs Internalities: A Model of Political Integration. *Journal of Public Economics* 68(2), 251–268.

Gordon, N. – Knight, B. (2009): A Spatial Merger Estimator with an Application to School District Consolidation. *Journal of Public Economics* 93(5–6), 752–765.

Imbens, G. (1992): The Efficient Method of Moments Estimator for Discrete Choice Models with Choice-Based Sampling. *Econometrica* 60(5), 1187–1214.

King, G. – Zeng, L. (2001): Logistic Regression in Rare Events Data. *Political Analysis* 9, 137–163.'

King, G. – Zeng, L. (2001): ReLogit: Rare Events Logistic Regression. *Journal of Statistical Software* 8(2), 137–163.

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

Manski, C. F. and Lerman, S. R. (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.
Moisio, A. – Loikkanen, H. A. – Oulasvirta, L. (2010): Public Services at the Local Level – The Finnish Way. VATT Policy Reports 2. Helsinki.
Musgrave, R.A. (1959): *The Theory of Public Finance: A Study in Public Economy*. New York: Mcgraw-Hill.
Oates, W. E. (1972): *Fiscal Federalism.* New York: Harcourt Brace Jovanovich.
Oates, W. E. (1999): An Essay on Fiscal Federalism. *Journal of Economic Literature*, Vol. XXXVII, 1120–1149.
Poirier, D. (1980): Partial Observability in Bivariate Probit Models. *Journal of Econometrics* 12(2), 209–217.
Tiebout, C. M. (1956): A Pure Theory of Local Expenditures. *Journal of Political Economy* 64(5), 416–424.
Weese, E. (2011): Political Mergers as Coalition Formation. Yale University, Economic Growth Center Discussion Paper 997.
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. – Rasche, F. (2006): FANMOD: a tool for fast network motif detection. *Bioinformatics* 22(9), 1152–1153.

Appendix. Additional figures and tables.

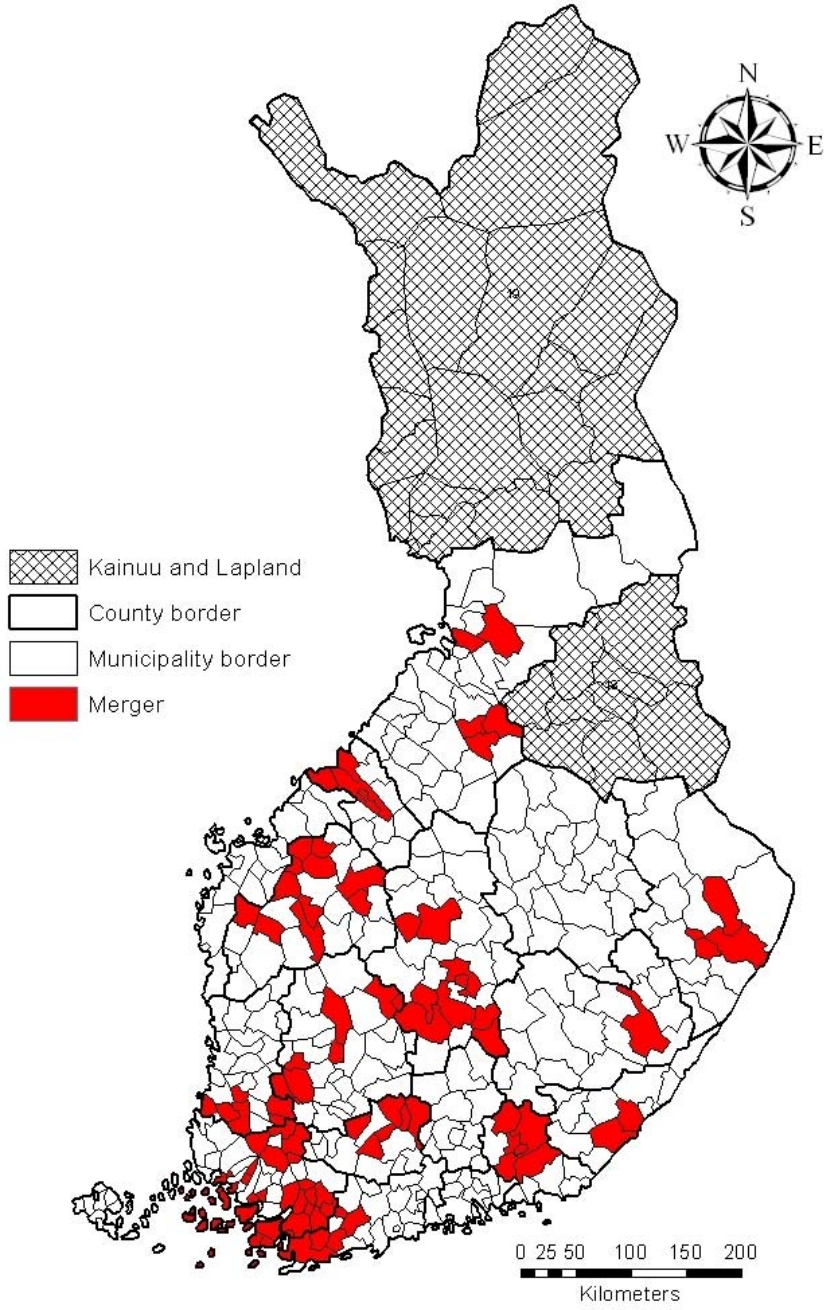

Figure A1. Map of municipality mergers in 2008–2009.

Table A1. Variable description.

Average coalition size	Average number of municipalities in a coalition.
Total population	Total population of the coalition.
Mean of taxable income (€)	Coalition level mean taxable income.
Total expenditures (€ per capita)	Total municipal expenditures per capita at coalition level.
Merger subsidy (€ per capita)	Central government merger subsidy per capita.
Cooperation	Dummy that equals 1 if all municipalities in the coalition organize basic health care through a joint authority, zero otherwise.
Median voter distance (km)	Median euclidian distance of voters to the centre of the largest municipality in the coalition. Centre corresponds to the location of the city or town hall. The voters of the largest municipality are excluded. Voters' location is based on 250 m x 250 m grids.
Language	Dummy that equals 1 if at least one municipality in the coalition is classified in a different language class than others in the coalition, zero otherwise. There are four different classes: unilingual Finnish, unilingual Swedish, bilingual with a Finnish speaking majority and bilingual with a Swedish speaking majority.
Population size heterogeneity	Maximum population - Minimum population.
Income heterogeneity	Maximum mean income - Minimum mean income.
Tax rate heterogeneity	Maximum municipal income tax rate - Minimum municipal income tax rate.
Expenditure heterogeneity	Maximum expenditure per capita - Minimum expenditure per capita.
Largest municipality vote share	(Number of votes given in the largest municipality in the coalition in the 2004 elections)/(All votes given in the coalition)
Herfindahl index	Standard Herfindahl index based on coalition level party shares. Coalition party shares are calculated by simply adding up all councilors in a coalition.
Same largest party	Dummy that equals 1 if all municipalities in the coalition have the same largest party, zero otherwise.
Centre Party share	The share of the Centre Party at coalition level.
Left-wing party share	The share of left-wing parties (Social Democrats and Left Alliance) at coalition level.
National Coalition Party share	The share of the National Coalition Party at coalition level.
Other party share	The share of other parties at coalition level.
Reduction in council size	Relative reduction in overall council seats if the merger goes through.
Party list size candidates	The share of 2004 candidates that would not fit into a party list if the merger goes through.
Party list size councilors	The share of 2004 councilors that would not fit into a party list if the merger goes through.
Municipal employees share council	The share of council members who are employed by the municipality sector.
Municipal employees share population	The share of municipal employees in the population.

BIS Department for Business Innovation & Skills

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

Spatial Economics Research Centre (SERC)
London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7852 3565
Fax: 020 7955 6848
Web: www.spatialeconomics.ac.uk

SERC is an independent research centre funded by the Economic and Social Research Council (ESRC), Department for Business Innovation and Skills (BIS) and the Welsh Assembly Government.