Size, Democracy, and the Economic Costs of Running the Political System

Jens Blom-Hansen Aarhus University
Kurt Houlberg KORA, Danish Institute for Local and Regional Government Research
Søren Serritzlew Aarhus University

The search for the optimal size of political systems is one of the most enduring in political thought. Given the validity of arguments for and against small units, one might expect variation in rearrangements of unit sizes. However, the reform trend is uniform: units, often at the local level, are amalgamated to harvest scale effects. The purpose of this article is to evaluate the argument on economies of scale in the economic costs of running political systems. Our testing ground is a recent Danish reform. It allows us to avoid endogeneity problems often facing researchers of size reforms. The reform was directed by the central government and constitutes an exogenous shock to 239 municipalities, whereas 32 municipalities were left untouched. We thus have a quasi-experiment with pre- and posttreatment observations for both an experiment and a control group. Our findings show that scale effects, measured as administrative costs per inhabitant, are considerable.

Is there an optimal size for political systems? This is one of the most enduring questions in political thought, and many thinkers have pointed to the dilemma that there may be a trade-off between democratic and economic concerns. In classical Greece, Plato and Aristotle stressed the desirability of small, autonomous entities where all citizens, for democratic purposes, could know each other, although the entities should also, for economic reasons, be large enough to support themselves.

In more recent times, political thinkers have also seen considerable democratic virtues in small jurisdictions. They argue that smallness facilitates citizens' participation in politics, enhances their trust in their own political competence, and breeds civic consensus. It makes politics less abstract and increases politicians' responsiveness to citizen views. It spreads political power, furthers control over government, increases political accountability, and facilitates exit-based empowerment of citizens. However, there are also important democratic arguments against small jurisdictions. In large jurisdictions, there is allegedly more diversity in beliefs and values, and politics becomes more competitive and professionalized. Large units have greater system capacity and can supply a greater range of public services. They are less vulnerable to the influence of local business and have more serious media coverage of local politics. They also have more organizational activity and thus more community groups, interest organizations, and political parties (Dahl 1967; Dahl and Tufte 1973, 4–16; Lassen and Serritzlew 2011; Lewis 2011; Newton 1982; Sharpe 1970; Treisman 2007, 1–19; Warren 2011).

From an economic perspective, there are important arguments both for and against small jurisdictions. For many years, economists have argued that small entities create the potential for welfare gains because public services can be better tailored to local preferences (Oates 1972, 31–63) or because citizens can move to localities that offer the ideal tax-service package (Tiebout 1956). At the state level, a similar logic applies. Smaller nations enjoy the benefit of a more homogeneous population, with more homogeneous preferences (Alesina and Spolaore 2005). However, an equally long-lived counterargument holds that large jurisdictions are more cost-effective due to economies of scale in the production of many public functions (Hirsch 1959). Size also entails benefits for nations. For example, a larger population allows for cheaper

Jens Blom-Hansen is Professor, Department of Political Science, Aarhus University, Bartholins Allé 7, 8000 Aarhus C, Denmark (jbh@ps.au.dk). Kurt Houlberg is Director of Research, KORA, Danish Institute for Local and Regional Government Research, Købmagergade 22, 1150 Copenhagen, Denmark (kuho@kora.dk). Søren Serritzlew is Professor, Department of Political Science, Aarhus University, Bartholins Allé 7, 8000 Aarhus C, Denmark (soren@ps.au.dk).

The authors would like to thank David Dreyer Lassen, Ulrik Hvidman, Peter Bjerre Mortensen, Kim Mannemar Sønderskov, and four anonymous referees for comments on various drafts. The study was partly funded by a grant from The Danish Council for Independent Research. Replication files can be found in the AJPS Data Archive on Dataverse (http://dvn.iq.harvard.edu/dvn/dv/ajps).

American Journal of Political Science, Vol. 58, No. 4, October 2014, Pp. 790–803

DOI: 10.1111/ajps.12096

provision of many public goods, as well as for larger markets that are not limited by political borders (Alesina and Spolaore 2005).

While the size of nations changes quite rarely, lower tiers of government are sometimes redesigned (Treisman 2007). In the past decades, municipal boundaries have been redrawn, in comprehensive or limited ways, in a large number of countries. Given the validity of the arguments both for and against small jurisdictions, one might expect a great deal of variation in the direction and intention of these reforms. However, the reform trend has been uniform: to reduce costs by making local governments larger. The agenda has been one of municipal amalgamations as an instrument to harvest scale effects (Baldersheim and Rose 2010; Caulfield and Larsen 2002; King and Ma 2000; Sancton 2000; Vetter and Kersting 2003a, 2003b).

In other words, to practical reformers, scale effects in local service production seem so important that they trump other arguments on jurisdiction size. But can this particular argument on the optimal size of local governments really bear this burden? The purpose of this article is to evaluate the argument on economies of scale in local government. We first review the theory behind the proposition and discuss the existing evidence, which turns out to be mixed. We argue that this state of affairs may be due to a number of methodological challenges facing cross-sectional investigations of scale effects. We then discuss how these challenges can be met by studying a recent reform of Danish municipalities. In the period 2005–11, 239 municipalities were amalgamated into 66 new entities, whereas 32 municipalities were left untouched. Since the reform was directed by the central government, it constitutes a (largely, as we shall argue) exogenous shock to the local government system. We thus have a quasi-experiment with pre- and posttreatment observations for both an experiment and a control group. This is a strong design for the estimation of the causal effect of changes in jurisdiction size.

Our findings show that scale effects are considerable. Increasing the size of the jurisdiction makes it possible to run the political system with fewer economic costs. However, there may be a price to pay in terms of reduced welfare gains and less democracy. In the conclusion, we return to this issue and discuss the dilemmas involved in the search for the optimal jurisdiction size.

Economies of Scale in Local Government: Theory and Evidence

Economies of scale refer to how output responds to variation in input. If there are increasing returns to scale (positive scale effects), large entities can produce with lower unit costs than small ones. Most production processes first exhibit increasing returns to scale, then constant returns, and finally decreasing returns to scale. The normal average cost curve is thus U-shaped. This proposition holds for both private and public production (Boyne 1995; Hirsch 1959; Sawyer 1991, 47–49).

There are two main reasons why production processes initially exhibit increasing returns to scale. First, some production costs are constant. Even small productions need buildings and machinery, and even a small local jurisdiction needs a mayor and a mayoral office. As production grows, these fixed costs can be spread over increasing units of production. Second, increasing production allows for more division of labor into more specialized functions. As scale increases, benefits of specialization can thus be reaped. However, after a certain level of production is reached, decreasing returns to scale may begin to emerge and unit costs to rise. The main source of this effect is problems of communication and coordination. As production grows, problems of transmitting accurate information through increasing layers of management become more and more serious. Controlling the decisions that are put into effect also becomes more difficult as scale increases. In short, large production processes suffer from increasing management costs.

The expected U-shaped cost curve suggests that there is an optimal size of local government. However, the theory says nothing about the exact location of the trough of the curve. This is an empirical question depending on production technologies, relative input costs, and so on. For this reason, the location of the trough is likely to vary across different public outputs. For instance, the optimal size of government in the road area may be different from the school area. This means that finding the optimal size for a jurisdiction requires either splitting a multipurpose unit into single-purpose units or averaging across several different public outputs (Ostrom 1972).

When focusing on multipurpose local governments, it would, for two reasons, be a futile task to try to estimate an optimal size in general. First, the policies provided by the local tier of government vary across countries. In some countries, the local level is involved in limited property-oriented services. As Boadway and Shah (2009, 276) put it, Australian local government is responsible for only "road and rubbish," whereas the Nordic local governments provide a wide range of services, from rubbish to social security and many comprehensive welfare services. However, common to all types of local governments is that they must be governed. In this sense, at the core of local government is the cost of running the political system. Any local government is a political system that must

be run, and this involves administrative costs. Around the world, any other task or public output can be—and is—placed in different tiers of government. Therefore, we focus on the economic costs of ruling, that is, of running the local government system or the administrative costs. Here, it makes sense to ask whether large jurisdictions can operate with fewer costs than small ones.

The second issue concerns the distinction between plant-level and firm-level returns to scale. Studies focusing on population when estimating scale effects in local government assume that scale effects are a function of the size of the local authority as a whole, not the size of individual schools, child care centers, residential homes for the elderly, and other municipal institutions (Boyne 1995, 220; Sawyer 1991, 50–51). If scale effects are possible at both levels (Scherer and Ross 1990), it is important to disentangle them since the size of service-providing institutions can vary for authorities with identical populations. In a Danish setting, it has been demonstrated that there are scale effects in the primary school system at both the level of the municipality (Houlberg 2000, 27–33) and the individual school (Blom-Hansen 2004). Our contention is that in most policy areas, scale effects in local government are a plant-level, not a firm-level, phenomenon. For instance, in the school area, most expenditures (i.e., wages to teachers, teaching material, buildings) are spent at the level of the individual schools, so the most relevant scale question in this area is whether small schools are more expensive than large ones. This is another reason why we focus on administrative costs, which are (mostly, as we shall argue) a firm-level phenomenon.

Many studies have investigated scale effects in local government, and Table 1 reports the results of studies that summarize findings on this question. The table is thus a survey of survey studies, which at first sight is disappointing news for believers in economies of scale in local government. There is, at best, mixed evidence on this theory. No survey study has found unequivocal evidence in favor of positive scale effects. However, mixed evidence should not necessarily come as a surprise in this area. Cross-sectional analyses of scale effects face a serious problem of endogeneity. In the next section we discuss this problem and how quasi-experiments can help alleviate it.

Endogeneity and a Quasi-Experiment

The basic problem in studying the effect of jurisdiction size is that the size of local government entities is typically not exogenously determined. Size rather reflects active responses to problems. This implies that size is not only an isolated variable with effects on the dependent variable, but also a consequence of factors related to the dependent variable. Hence, traditional cross-sectional studies face potentially serious problems of endogeneity, which can arise in several ways. A jurisdiction experiencing economic problems may opt for an amalgamation in order to ease these problems. Or a jurisdiction may want an amalgamation in order to raise quality levels and thus increase costs. To the analyst studying the relationship between size and costs, these situations constitute problems of reverse causality, or policy endogeneity (Besley and Case 2000). They may bias studies in any direction, but perhaps most likely toward null findings since the reverse effects may cancel each other out. Most studies of scale effects in local government are cross-sectional (e.g., Andrews and Boyne 2009). Their estimates may therefore be biased due to past decisions of local policy makers to amalgamate in order to solve problems. Cross-time variation in jurisdiction size is rare, and studies making use of such variation (e.g., Benton and Gamble 1984) are vulnerable to selection bias due to the voluntary nature of most amalgamations. The solution is to study exogenously induced changes in jurisdiction size. This is what the Danish municipal reform in 2007 makes possible (cf. our discussion of this reform below).

If scale effects exist in local government, we should, all things equal, observe a negative correlation between municipality size and per capita administrative costs. However, there are many differences between small and large jurisdictions, and omitted variables and simultaneity are likely to create problems of endogeneity, which renders correlations in cross-sectional studies poor estimates of causal effects. Rather than computing the correlation between size and administrative costs, we therefore follow the logic of experimental methods, according to which random assignment of subjects to treatment and control groups is an effective solution to the problem. In the present case, this would involve randomly assigning municipalities to a treatment and a control group. The size of municipalities in the treatment group should then be increased, whereas the size of municipalities in the control group should remain constant. If municipalities are assigned randomly, which we do not attempt to do in this study, any systematic difference in administrative costs between the treatment group and the control group must be due to the change in jurisdiction size, not endogeneity, and can be interpreted as an estimate of the causal effect.

Although truly random assignment is rare in experiments that are not fully controlled by the researcher, such as laboratory experiments, assignment to treatment groups is sometimes exogenously determined. We then

TABLE 1 Survey Studies of Scale Effects in Local Government

	Surveyed Studies on Scale Effects	Main Conclusion of Survey
Ostrom (1972)	13 U.S. studies	No association between size and costs
Derksen (1988)	Unspecified number of Dutch amalgamation studies	No economic gains from amalgamations
Boyne (1995)	17 UK studies	Diseconomies of scale (unit costs rise with scale)
Martins (1995)	Unspecified number of European studies	Mixed results; no conclusive evidence on scale effects
Bish (2001)	Unspecified number of U.S. and Canadian studies	80% of local government activities do not possess economies of scale beyond relatively small municipalities
Byrnes and Dollery (2002)	32 UK, U.S., and Australian studies	Mixed results; no conclusive evidence on scale effects

have a quasi-experiment rather than a laboratory experiment or a natural experiment.[1] Quasi-experiments are "studies in which there is a transparent exogenous source of variation in the explanatory variables that determine the treatment assignment. [A quasi-experiment] induced by policy changes, government randomization, or other events may allow a researcher to obtain exogenous variation in the main explanatory variables. This occurrence is especially useful in situations in which estimates are ordinarily biased because of endogenous variation due to omitted variables or selection" (Meyer 1995, 151). This is exactly the core problem in identifying the effects of jurisdiction size on administrative costs in traditional cross-section analyses.

The Danish structural reform of 2007 constitutes a quasi-experiment that solves this problem.

It has been used before to estimate the causal effect of jurisdiction size on democracy (Lassen and Serritzlew 2011). Our aim is to use the same logic to estimate the causal effect of jurisdiction size on the other side of the classic trade-off: economies of scale. Before turning to the structural reform, we give a short introduction to Danish local government.

The Danish public sector has a regional tier with five units mainly responsible for health care, and a local tier with 98 municipalities responsible for the main components of the welfare state. The municipalities run welfare services such as the public schools, child care, and elderly care. Of the Danish gross domestic product (GDP) of EUR 237 billion (2010), EUR 58 billion is spent by municipalities, corresponding to about 42% of the total public expenditure. Municipal administration accounts for EUR 4 billion (2010). Hence, Danish municipalities are multipurpose units responsible for administrating large, expensive, politically sensitive, and important tasks. They are governed by city councils with between 9 and 31 members elected by citizens every four years. The mayor, elected by and among local councilors, is head of the administration. However, administrative costs are, just as spending on other purposes, set in the budget, which is enacted by the council.

The Danish structural reform of 2007 had three aspects. First, 14 counties were amalgamated into five regions. Second, tasks were transferred from counties to the state (e.g., youth education and larger roads) and to the municipalities (e.g., environmental regulation and specialized institutions for the disabled), and from municipalities to the state (e.g., tax assessment). In total, these changes were quite limited for the municipalities. The net transfer of tasks amounted to EUR 3.3 billion, a small share of total municipal expenditure. Third, and most importantly for our purposes, 271 municipalities[2] were by law amalgamated into 98.

The reform was quick and radical (Mouritzen 2010). Before 2002, the idea of a centrally imposed structural reform played no role in the debate, and when mentioned in a parliamentary committee debate, it was firmly rejected by the Minister of the Interior. However, in 2002, the government formed the Commission on

[1]There is not agreement in the literature about the exact difference between a natural experiment and a quasi-experiment. However, following Cook and Campbell (1979, 56), Dunning (2012, 15–21), and Sekhon and Titiunik (2012), we define a *natural experiment* by the random, or as if random, assignment of subjects to treatment, whereas the term *quasi-experiment* characterizes exogenous, but not random, assignment to treatment. According to this definition, our experiment is a quasi-experiment.

[2]Including *Ærøskøbing* and *Marstal*, which were amalgamated into *Ærø*, effective January 1, 2006.

Administrative Structure, with members appointed by the Minister of the Interior, three expert members, and representatives from municipalities and counties. In 2004, the commission published its report, which recommended amalgamations. In fewer than six months, a majority in the national parliament decided on a semi-voluntary amalgamation reform (Regeringen and Dansk Folkeparti 2004, 9–12). It stipulated that municipalities with fewer than 20,000 citizens should be merged with neighbors, and that amalgamations should create new municipalities of about 30,000 citizens. This was considered the minimum size for the provision of services of sufficient quality. Municipalities with fewer than 20,000 inhabitants could only avoid amalgamation if they entered into a cooperative arrangement with a large neighboring municipality about service provision. This proved very difficult in practice, and only five municipalities out of more than 200 with fewer than 20,000 inhabitants succeeded in avoiding amalgamation by making such a cooperative arrangement. The reform agreement explicitly stated that if this arrangement was not made, and if a small municipality nonetheless did not want to amalgamate, then the reform partners in Parliament would decide on the amalgamation. This threat of central intervention was effectuated in relation to three small municipalities that neither found amalgamation partners on their own nor made cooperative arrangements with neighbors (namely, the municipalities of *Farum*, *Holmsland*, and *Hvorslev*). While the impetus for amalgamation was thus clearly exogenous to the individual municipality, the precise choice of amalgamation partner and the exact size of the new amalgamated municipality were left to local decisions. The reform gave the old municipalities six months to settle the amalgamations. The key issue for our research question is whether administrative costs played any large role for the individual municipalities' choices in this process, since this would introduce an element of endogeneity in our design. However, qualitative and quantitative evidence suggest that administrative considerations were, at the most, secondary. Case studies of individual amalgamation processes suggest that several other considerations were important, including local identity, internal cohesion, the likely political party composition of the future municipality, homogeneity in service and wealth, and, not least, ambitions of becoming an influential player in the future municipal structure (Mouritzen 2006). A quantitative investigation of all amalgamations suggests that, in the end, preexisting patterns of social connectedness were decisive. A regression analysis of several potential determinants of the amalgamation decisions shows that only citizens' commuting patterns, in addition to the obvious size parameter, can predict the actual amalgamation decisions (Bhatti and Hansen 2011). In sum, we feel confident that considerations of administrative costs did not play a significant role in the actual amalgamation choices. In other words, the amalgamation reform constituted a largely exogenous shock to the municipalities' administrative systems.

Following the central reform agreement, local mergers were arranged by the municipalities over the next six months. As noted, the threat of central intervention was only effectuated in a few places. Very few small municipalities survived because the option of entering into a cooperative arrangement with a large neighboring municipality proved difficult in practice, since few large municipalities wanted these complicated arrangements. The end result was 98 new municipalities. Of these, 66 were results of amalgamations of 239 old municipalities, whereas 32 municipalities remained unchanged. The amalgamations took effect from January 1, 2007. Hence, the structural reform implied that a treatment group of 239 experienced an exogenously induced shock to their size, whereas a control group of 32 experienced no change in size apart from usual demographic changes.

Table 2 shows the change in size for the treatment and the control group. Municipalities in the control group experienced very little change in size (average size increased from 62,091 to 62,174). In the treatment group, changes were dramatic. In 2006, almost half the municipalities had 5,000–10,000 citizens compared to 30,000–50,000 after the reform.

We use these exogenously induced changes in size to study scale effects. Since this approach uses amalgamation reform in order to get leverage to address problems of endogeneity, it is important to consider how the effects of increases in size can be disentangled from the effects of reform itself. One possibility might be that a large reform in itself, not the change in size, could provide an opportunity for reorganization, which may then again lead to cost savings. However, the existence of a control group of 32 municipalities, which were not amalgamated but experienced other aspects of the municipal reform (cf. above), allows us to use a difference-in-difference (DiD) estimator to isolate the effect of amalgamations from the effect of reform itself.[3]

Although assignment to treatment and control groups was exogenously determined, and although register-based data for administrative costs are available for all municipalities, before and after the reform, the

[3]An ideal solution to this potential problem could be to, in addition to amalgamations, study economies of scale in municipalities splitting up. If scale effects exist, splits should produce higher per capita costs. However, the Danish reform did not result in splits.

TABLE 2 Size of Municipalities in Control Group and Treatment Group, before and after Reform (Percent)

	Control Group		Treatment Group	
	Prereform	Postreform	Prereform	Postreform
Under 5,000	9	9	5	2
5,001–10,000	0	0	47	0
10,001–20,000	6	6	31	2
20,001–30,000	28	28	7	14
30,001–50,000	31	31	5	44
50,001–100,000	16	16	3	35
More than 100,000	9	9	0	5
Total	100	100	100	100
N	32	32	239	66

effect of increasing jurisdiction size cannot be readily estimated because the assignment to treatment and control groups was not randomized. We therefore, as in Lassen and Serritzlew (2011, 245–47), use the DiD approach. Let the administrative costs per capita for municipality i be Y_i^1 in the treated case and Y_i^0 in the untreated case. We are interested in estimating $\Delta_i = Y_i^1 - Y_i^0$. The problem, of course, is that municipality i cannot simultaneously be both treated and untreated. Hence, either Y_i^1 or Y_i^0 will be missing. Instead, we can observe $Y_i^{TG=1,T=1}$ and $Y_i^{TG=1,T=0}$, where $T = 0$ and $T = 1$ indicate that i is observed in 2006 and in 2007, respectively, and $TG = 1$ indicates that i belongs to the treatment group. We can then estimate $Y_i^{TG=1,T=1} - Y_i^{TG=1,T=0}$, which is the combined effect of the treatment and other changes as time has passed from 2006 to 2007.

We use the DiD approach to estimate the change in administrative costs in control and treatment municipalities, before and after the reform. The basic logic is that we can estimate the combined effect of treatment and time by calculating the difference between costs in 2007 and 2006 for municipalities in the treatment group ($Y_i^{TG=1,T=1} - Y_i^{TG=1,T=0}$, the first difference), and the effect of time by calculating the difference between costs in 2007 and 2006 for the control municipalities ($Y_i^{TG=0,T=1} - Y_i^{TG=0,T=0}$, the second difference). The DiD estimator is the difference between these two differences: the estimate of the combined effect of treatment and time, minus the estimated effect of time. We end up with an estimate of the effect of the treatment. In other words, the DiD estimator is an estimate of the treatment effect for the treated (ATT). Formally, we obtain this estimate from the following regression analysis, where T is a dummy variable (equal to 1 postreform) and TG is as defined above:

$$Y_i = \alpha + \beta_1 TG_i + \beta_2 T_i + \beta_3 TG_i \times T_i + \varepsilon_i. \quad (1)$$

The estimate of the effect of size (the DiD estimator) is then

$$\begin{aligned}&\left(\bar{Y}_i^{TG=1,T=1} - \bar{Y}_i^{TG=1,T=0}\right) - \left(\bar{Y}_i^{TG=0,T=1} - \bar{Y}_i^{TG=0,T=0}\right) \\ &= \alpha + \beta_1 + \beta_2 + \beta_3 - (\alpha + \beta_1) - ((\alpha + \beta_2) - \alpha) = \beta_3.\end{aligned} \quad (2)$$

Hence, we estimate the effect of size on costs based on prereform and postreform observations of administrative costs for municipaties in the treatment group, which experienced marked changes in size, and in the control group, which did not. The model can easily be expanded in two ways. First, T (indicating pre-/postreform) can be substituted with dummy variables for several time periods pre-/postreform. This is advantageous because amalgamations may be costly to implement. Scale effects may consequently not be visible in the first year after the reform. We therefore expand the model with dummy variables T_{2006}–T_{2011} and corresponding interaction terms $TG_i \times T_{2006} - TG_i \times T_{2011}$. The interaction terms produce six DiD estimates, β_3^{2006}– β_3^{2011}. The first, β_3^{2006}, is an estimate of economies of scale just before the reform. We expect this effect to be zero. β_3^{2007}–β_3^{2011} are estimates of scale effects in the five years following the reform. If the costs of running a political system decrease with size, these estimates will be negative. If the scale effects appear gradually, the estimates should grow from 2007 to 2011. Second, the model can be expanded by a set of time-varying control variables to account for possible differences between the treatment and control groups over time (Wooldridge 2009, 454). By including these controls, we can minimize potential bias resulting from systematic differences in changes over time between the treatment

and control groups. We return to this below. In the next section, we discuss issues of measurement and data. We then turn to the empirical analysis.

Methods and Data

Measuring scale effects is not straightforward. Before turning to the data set, we discuss three issues: how cost conditions can be estimated, the distinction between scale and scope effects, and the measurement of output. Cost conditions in local government can be estimated in different ways. The issue here is how to identify the scale at which local governments produce the most at the lowest costs. In the private sector, three scale estimation methods are used, which to varying degrees are relevant for local governments (Bish 2001, 12–15; Boyne 1995, 215–19; Sawyer 1991, 51–61). First, the engineering method breaks down the production process into its component parts and estimates the cost condition for each part. This method requires detailed knowledge of input factors and their interaction. It may be useful for the analysis of specific local government functions, such as road services (Deller, Chicoine, and Walzer 1988) or cleaning of schools (Christoffersen and Paldam 2003), and may therefore have considerable potential for studying the optimal size of single-purpose authorities. But if the objective is to estimate the optimal size of multipurpose authorities, it is impractical since the knowledge of input factors that are required to estimate production functions is rarely available. Second, the survivor technique observes the size of successful firms in competitive markets, which is then taken as an indication of the optimal scale of production. This method is largely irrelevant for local government studies since there is no guarantee that inefficient (too small or too large) authorities will go out of business. Finally, the statistical estimation of cost curves uses observational data to estimate the relationship between size and average aggregate costs. In practice, this is the only method open to the study of multipurpose authorities, and therefore it is the most frequently used in the literature. It is also our choice of method, but one drawback should immediately be noted: it assumes that the studied authorities produce efficiently, or that any deviance from technical efficiency is distributed randomly. This may be a "heroic" assumption, and we will return to it in the conclusion.

The second issue is how to keep scale effects separate from scope effects. While scale effects refer to the relationship between input and output for a given function, or set of functions, scope effects refer to the efficiency potential of varying the types of outputs (Dollery and Fleming 2006). Since it is likely that both types of effects will result in different cost curves for authorities that vary in both size and functions, it is important to disentangle their relative impact. We study Danish municipalities before and after the amalgamation reform in 2007. As mentioned, the reform also resulted in new functions for the municipalities and thus potentially gave rise to scope effects. However, as already explained, a number of municipalities were not amalgamated and thus serve as a control group for the estimation of scale effects. Hence, the difference-in-difference design isolates the effect of changes in jurisdiction size.

The third issue is how to estimate outputs, units, and unit costs in a local government setting. Following Hirsch (1959), most studies use population as a measure of scale and as a proxy for output, and measure unit costs as expenditure per capita. However, as noted by Tiebout (1960) and Boyne (1995, 219–21), this is less than ideal if the ambition is to measure technical efficiency. Expenditure per capita is not the same as the economic concept of unit costs. However, given the lack of output measures for the sum of local government activities, and hence the difficulty of using the efficiency concept at this level and thus disentangling quality from costs, we are content to measure costs as expenditure per capita. Seen from the taxpayers' perspective, this is probably the most relevant cost measure.

In sum, we investigate scale effects in local government by focusing on administrative costs per capita. We employ a design that makes it possible to avoid entangling scale and scope effects, to avoid mixing up plant- and firm-level effects, and to avoid the endogeneity problems often found in the literature. We now explain the data set in more detail.

The Danish structural reform took effect in 2007. Our data cover 2005–2011, that is, two years before the reform and five years after. To compare municipalities before and after the reform, we impose the post-2007 structure on the pre-2007 structure. This means that we aggregate prereform municipalities that were amalgamated in 2007 to their postreform size.[4] We exclude three municipalities with prereform two-tier status as both county and municipality and thus end up with a data set containing 95 cases across seven years, a total of 665 observations.[5]

[4]In the few cases where parts of an old municipality were amalgamated with different new municipalities, the figures for the old municipalities are split according to the share of the population going to each new municipality.

[5]The multivariate analysis is based on 657 observations, since data for one of the variables are missing for one municipality in one year and we exclude an outlying municipality in all seven years.

Our measure of the impact of the amalgamations is the DiD estimator. As explained in the previous section, this is an interaction term between treatment (i.e., amalgamation) and time. Since the exact timing of the amalgamation effect is theoretically indeterminate, we include interaction terms for all postreform years. This allows us to investigate whether the effect occurs with a time lag.

The dependent variable is net current administrative costs per capita measured in DKK at the 2011 price level (1 Euro equals approximately 7.4 DKK). Administrative costs include a number of accounts in the accounts of the Danish municipalities (see the supporting information, Table S1, for details) encompassing mainly wages for administrative personnel, emolument for politicians, maintenance of buildings, purchasing of administrative utensils, insurance, auditing, and so forth. Wages account for approximately 75% of the administrative costs, whereas emolument for politicians accounts for less than 2.5%. We include only current expenditures, since capital expenditures in Denmark are fully accounted in the year of investment. Our measure of current expenditures does not include depreciations, as these are not accounted in a systematic and comparable way. We calculate the costs net of administrative income, as some municipalities run administrative functions for other municipalities or joint intermunicipal partnerships.

Using this measure of administrative costs presents three challenges. First, municipalities' accounts contain self-reported figures, so there may be some variation in accounting practices across years and municipalities. But since accounting is done according to a relatively detailed uniform system specified by the central government and controlled by independent auditors, potential variation is not likely to be systematic. Second, since Danish municipalities' functions are not constant over the period of investigation, we have adjusted the administrative costs for functional changes. All figures are corrected to indicate functions in 2011.[6] Third, administrative expenditure at the town hall level in each municipality is registered on a separate account, but administrative expenditure at the level of institutions (schools, libraries, day care centers, etc.) is not. Since the municipalities decentralize their administrative functions to their institutions to different degrees, we include a control for these differences. In practice, since decentralized administrative costs are not identifiable in the municipal accounts, we include a measure for administrative decentralization of the staff as a control variable in our regression analyses. See Table S1 in the supporting information for the exact construction of this measure.

Municipal administrative costs are influenced by factors other than jurisdiction size. And since assignment of municipalities to treatment and control groups is not randomized, we include standard control variables known from previous analyses of Danish municipalities (Houlberg 2000; Houlberg 2011; Mouritzen 1999). First, we include four indicators for administrative spending needs: dispersed settlements (travel time index), municipalities on islands, and complicated social problems (social housing, single-parent children). Second, an indicator of fiscal pressure controls for variations in economic potential among the municipalities. Third, as some administrative functions are decentralized to individual schools, child care centers, and so on, we control for differences in decentralization of administrative staff. Descriptive statistics for dependent variables and control variables are provided in the appendix, for amalgamated and nonamalgamated municipalities, respectively.

Since our data have a panel structure, ordinary least squares (OLS) regression is not likely to produce unbiased results because of problems related to autocorrelation and heteroskedasticity. The literature suggests various remedies to these problems. We employ three often recommended solutions to investigate the robustness of our results. First, we run OLS regressions with cluster-corrected standard errors (clustered at the level of the individual municipality; Williams 2000). This is our basic model. Second, since annual changes are less likely to suffer from serial autocorrelation than levels, we run the same regressions, but with all time-variant variables changed into first differences. Third, we analyze a model with panel-corrected standard errors, including a lagged dependent variable, as recommended by Beck and Katz (1995).

[6]To be more specific, our administrative expenditure data have been adjusted for the administrative costs related to the functions the municipalities received and lost in the 2005 tax assessment reform, the 2007 structural reform, and the 2009 employment reform. These adjustments are based on the estimates of the administrative consequences of these reforms agreed by the central government and the local government association and formally transformed to positive and negative economic compensations to municipalities in the annual national act on grants for the municipalities. See Table S2 in the supporting information.

Results

In this section, we first report the results. We consistently find statistically significant scale effects. In all models, increases in jurisdiction size lead to lower administrative costs per capita. An assessment of the substantive significance of the findings follows. We estimate the average

scale effect to 10% of administrative costs, a quite dramatic effect.

Model 1 of Table 3 is an OLS regression with standard errors clustered at the level of the individual municipality. The dependent variable is the level of administrative costs per capita from 2005 to 2011. Hence, we have, for each municipality, two pretreatment observations (in 2005 and 2006) and five posttreatment observations (in 2007, 2008, 2009, 2010, and 2011). We estimate treatment effects for each year (i.e., the change compared to 2005 in administrative costs), following the logic in the DiD approach.

The dummy variable *Amalgamated* indicates whether the municipality belongs to the treatment group.[7] Hence, this entry in the table corresponds to β_1 in Equation (1). It is the estimate of the prereform (2005) difference between the treatment and control groups. If municipalities were, directly or indirectly, selected for treatment based on their administrative costs per capita, the effect of this dummy variable would be large. It turns out that the effect is not significantly different from zero. This indicates that the selection of municipalities for treatment is in fact, as expected, exogenous to administrative costs per capita. Since we have two prereform observations, we also estimate, by the interaction term *Amalgamated* × *2006*, the difference between treatment and control municipalities in 2006, just before the reform was implemented. Again we find, as expected, that the difference is small and statistically insignificant.

We also estimate the general time trend, not caused by changes in size, in administrative costs from 2005 to 2011. This is done by the dummy variables *2006*, *2007*, *2008*, *2009*, *2010*, and *2011*, corresponding to β_2 in Equation (1). It turns out that administrative costs have generally increased in the period. But in 2010, the increase was halted because the fiscal crisis hit Danish municipalities and forced them to cut administrative costs.

Finally, as described above, we include six control variables. First, *Dispersal of Settlement* shows that thinly populated municipalities, measured by an index of the average travel time in each municipality, tend to have higher administrative costs. This could be an effect of the need for more civil service centers in more dispersed municipalities (see Bhatti, Olsen, and Pedersen 2011).

The dummy variable *Small Island* shows that five small islands have extraordinary administrative costs. *Social Housing* and *Children in One-Parent Families* are two indicators of social problems. As expected, both are associated with higher spending. In accordance with our expectations, we also find that municipalities facing *Fiscal Pressure* have lower administrative costs. Finally, we control for differences in *Decentralization of Administrative Staff*. As expected, decentralization leads to lower central administrative costs.

The effect of size on administrative costs per capita can now be estimated. The variables *Amalgamated* × *2007*, *Amalgamated* × *2008*, *Amalgamated* × *2009*, *Amalgamated* × *2010*, and *Amalgamated* × *2011* are interaction terms indicating whether a municipality is treated and observed postreform. Hence, they are, as shown in Equation (2), DiD estimators. The table entries correspond to β_3 in Equation (1). In 2007, just after the reform, the effect is very small and statistically insignificant. This indicates that scale effects were not reaped immediately after the amalgamations. However, in 2008, the estimated effect is negative (but not statistically significant), and in 2009, 2010, and 2011, the negative estimated effect is increasing, large, and statistically significant. In 2011, the parameter estimate is –540, indicating that treatment status leads to, on average, DKK 540 per capita less spending on administrative costs. We conclude that size has a causal effect on administrative costs per capita. Scale effects do exist, but they do not appear until two to three years after amalgamations.

Models 2 and 3 are two other specifications, included as robustness checks. In Model 2, we estimate the effects of the same variables in a model with annual change in administrative costs as the dependent variable. If scale effects exist, we should be able to observe this not only in the level of administrative costs, but also in how administrative costs change after the reform. Since we use changes in spending as the dependent variable, and we do not have data for 2004, we have only one prereform observation. Again, the DiD estimators show that size has a causal effect on administrative spending. According to the estimates, administrative costs decrease in treatment municipalities after the reform. The effect is not statistically significant in 2007, just after the reform. In 2008 and 2009, the effect is quite strong and statistically significant. In 2010, the effect is still negative, indicating that municipalities are still reaping scale effects. The effect is, however, smaller and statistically insignificant. In 2011, the effect is approaching zero and scale effects seem to be exhausted. This indicates that the scale effects are realized mainly in the second and third year following amalgamation. It should be noted that administrative costs in 2006 related to amalgamation preparation are included in the reference year of this model (change in administrative costs from 2005 to 2006),

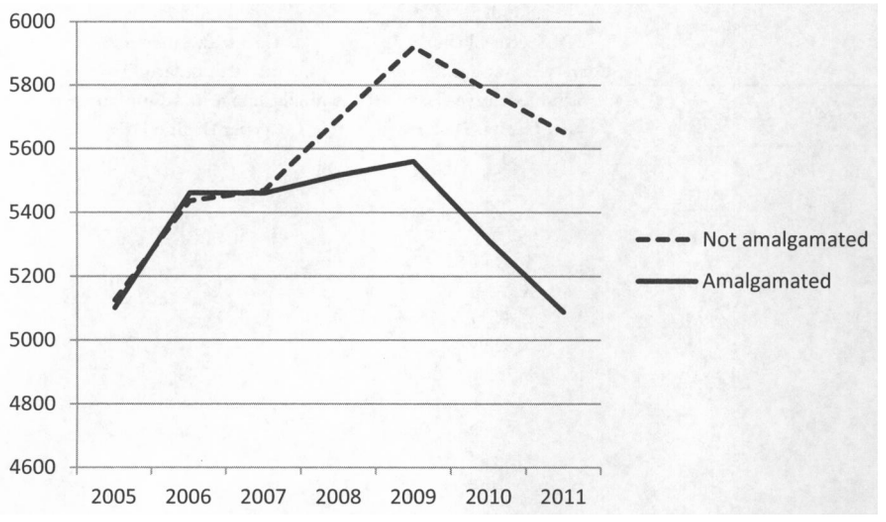
[7] An outlying municipality in this and other specifications was excluded due to deviant accounting and unpredictably low administrative costs (Houlberg 2011). In Model 1, Cook's D for this municipality is 0.04–0.05 for all years, approximately seven times higher than the value of 1/N, which is considered high (Kohler and Kreuter 2005, 211).

TABLE 3 Regression Analysis of Municipal Administrative Costs per Capita, 2005–2011

Estimation Method	Model 1 OLS with Cluster-Corrected Standard Errors	Model 2 OLS with Cluster-Corrected Standard Errors	Model 3 OLS with Panel-Corrected Standard Errors
Dependent Variable	**Administrative Costs per Capita (DKK)**	**Annual Change in Administrative Costs per Capita (DKK)**	**Administrative Costs per Capita (DKK)**
Administrative costs per capita (lag)	—	—	0.80***
	—	—	(0.09)
Amalgamated (dummy)	−23.79	79.41	3.26
	(226.28)	(74.08)	(30.37)
DiD Estimators			
Amalgamated × 2006	48.11	—	—
	(84.99)	—	—
Amalgamated × 2007	15.83	−65.49	−32.94***
	(148.08)	(167.09)	(4.73)
Amalgamated × 2008	−156.08	−277.87**	−236.32***
	(188.63)	(112.03)	(5.22)
Amalgamated × 2009	−336.95*	−315.44***	−316.83***
	(195.09)	(107.47)	(11.78)
Amalgamated × 2010	−445.98**	−140.07	−199.79***
	(208.64)	(128.98)	(34.23)
Amalgamated × 2011	−540.21***	−18.52	−94.45*
	(196.16)	(124.01)	(51.34)
Year Dummies			
2006	311.34***	—	—
	(84.80)	—	—
2007	343.56***	−107.77	1.39
	(128.30)	(183.23)	(5.94)
2008	570.53***	27.52	140.03***
	(181.20)	(103.71)	(15.59)
2009	794.50***	101.96	218.14***
	(179.44)	(92.21)	(32.67)
2010	656.29***	−206.16	2.24
	(190.64)	(129.31)	(44.26)
2011	526.22***	−337.04***	−170.24***
	(170.30)	(121.17)	(45.83)
Control Variables			
Dispersal of Settlement (for Model 2: Δ)	168.23***	410.20*	57.97***
	(52.02)	(225.14)	(21.60)
Small Island	1,713.31***	—	289.08
	(411.95)	—	(257.03)
Social Housing (for Model2: Δ)	25.47***	−6.97	5.76*
	(7.55)	(24.94)	(3.31)
Children in One-Parent Families (for Model 2: Δ)	52.76**	8.20	13.56**
	(20.19)	(62.77)	(5.71)
Fiscal Pressure (for Model 2: Δ)	−42.67***	2.57	−5.31
	(11.95)	(6.97)	(4.78)
Decentr.of Administrative Staff (for Model 2: Δ)	−55.10***	−21.95*	−15.83***
	(11.29)	(12.03)	(5.43)
Constant	10,095.76***	132.80	1,924.13**
	(1,277.62)	(87.12)	(942.13)
Observations	657	562	563
Adj. R-squared	0.640	0.076	0.882
Number of municipalities	94	94	94

Note: Variance inflation factor (VIF) for DiD estimators and year dummies is 5.7–7.9, 5.1–6.6, and 5.4–6.8 for the three models, respectively. For all other variables, VIF is less than 2.6, 2.0, and 2.8, respectively.
***p < .01, **p < .05, *p < .1.

FIGURE 1 Predicted Administrative Spending (DKK per Capita)



Note: The figure shows the predicted administrative spending per capita for the amalgamated and nonamalgamated municipalities. The predicted values are computed from Model 1 in Table 3. All discrete control variables are set to sample means, whereas the dummy variable *Small Island* is set to zero.

making the conditions for realizing scale effects in amalgamated municipalities more favorable in Model 2 than in Model 1.

Finally, in Model 3, we estimate a model with panel-corrected standard errors, with the level of administrative costs per capita as the dependent variable, and with a lagged dependent variable, as recommended by Beck and Katz (1995). As expected, previous administrative costs are a powerful predictor of present administrative costs. Otherwise, the model confirms the conclusions. The DiD estimators show that size does have an effect on administrative costs.

An alternative approach to statistical control for confounding variables is to adjust for prereform observable differences by matching municipalities in the treatment group with similar municipalities in the control group. This approach has some shortcomings in this context, as it does not allow us to account for postreform changes in control variables, as we do in the OLS-based analyses, and since we cannot estimate separate effects for each of the postreform years. However, since it is possible to obtain good balance between the treatment and control groups, the approach is another useful robustness check. The results from matching are consistent with the OLS-based results: administrative costs tend to be reduced for treated municipalities.[8]

We now turn to assessing the substantial significance of the effects. Figure 1 shows the predicted administrative spending per capita as a function of time for the treatment and control groups. The predicted values are computed from the estimates in Model 1, with all control variables set to the sample means. The dotted line

[8]The results are summarized here; full results are available from the authors upon request. We use standard propensity score matching with single nearest neighbor in a data set of the 271 prereform municipalities (excluding three prereform municipalities, which were split between two or more postreform municipalities). We predict treatment status by prereform number of citizens, geographical location (inside or outside of prereform county of Copenhagen), and urbanization (measured as share of citizens in prereform municipality living in a town > 10,000 citizens). We find, as expected, that these factors significantly correlate with treatment status in the unmatched sample. In the matched sample, the predictors are insignificant, pseudo-R^2 is 0.003, and the LR-test for significance of the model is insignificant, with $p = .540$. We can, in other words, obtain balance between the control and treatment group by matching. We use the observed administrative costs per capita for 2011 as the equilibrium results and estimate the difference between administrative costs in 2011 and 2005. We find, using psmatch2 in Stata, a difference in growth in administrative costs in the matched sample of −2,881. The estimate is statistically significant, with $p < .001$. It is debated whether standard errors should be estimated with bootstrapping. With this procedure, the estimate is statistically significant, with $p = .01$. The estimate is considerably larger than in the OLS-based model. This is likely to be a result of repeated matching with a single untreated municipality with a very high propensity score. This municipality is, in the matching framework, weighted heavily since it is comparable to treated municipalities. The problem can be reduced by using two or more neighbors, but at a cost, since these models fail to produce perfect balance. With two neighbors, the estimate of the difference between treatment and control groups is −1,889, with three −1,688, and with four nearest neighbors, it is −1,501. All estimates are statistically significant at the .01 level.

represents spending in control municipalities, the solid line in amalgamated municipalities. The growing difference between the two lines shows that administrative spending becomes markedly lower for treated municipalities. In 2007, immediately after the reform, treatment municipalities spent 0.1% less than municipalities in the control group. In 2008, the savings amounted to 3.2% of average administrative costs in the municipalities; in 2009, 6.4%; in 2010, 8.4%; and finally, in 2011, 10.4%. This is far from a trivial amount.

Conclusion

The investigation of scale effects in local government involves both endogeneity and selection problems. The size of jurisdictions is typically changed to solve specific problems, such as high costs of running the system in small jurisdictions. This creates policy endogeneity and makes estimates from cross-sectional data biased. Time-series data are often plagued by selection problems because most amalgamations are voluntary. Identifying causal effects requires data with exogenous changes in jurisdiction size. This article provides an analysis of such a data set. In 2007, 239 Danish municipalities were amalgamated into 66 new entities, whereas 32 municipalities were left untouched. Since the reform was directed by the central government, it constitutes an exogenous shock to the local government system. To researchers, the reform represents a quasi-experiment.

Using a difference-in-difference design to isolate the effect of size changes from general cost trends and other differences between treatment and control groups and to trace the effect over time, we find that increases in jurisdiction size lead to lower administrative costs per capita. The savings do not occur immediately—only after two to three years. But then they are substantial. Our analyses show that the average effect for treated municipalities equals 10% of total administrative costs. Although substantial, this effect is probably a conservative estimate of the true cost-saving potential. As discussed in the article, we had to investigate scale effects by "the statistical method," that is, to estimate observed cost curves. This method assumes that observed production is efficient, or that any deviation from efficiency is distributed randomly. We can observe that the amalgamated municipalities reduce costs by 10%, but we do not know whether the amalgamations made even larger cost savings possible. Yet, it is not unreasonable to speculate that some of the cost-saving potential was used to grease the wheels of the amalgamation process. However, this issue affects virtually any experiment where some decisional elements serve as the causal mechanism (i.e., most social science experiments). What we measured is what actually happened—and, by extension, is likely to happen in similar contexts—not what could have happened had everything gone exactly according to some efficiency ideal.[9]

Hence, evidence from the Danish municipal reform strongly suggests that the economic costs of running the political system decrease with jurisdiction size. It is important to consider the generalizability of these results. Danish municipalities are multipurpose units with a comparatively large portfolio of policy tasks, and with a population that typically is larger than 5,000 and smaller than 100,000. The results should only be generalized with caution to other types of systems. For instance, we cannot know, based on this reform, whether decreasing returns to scale set in for local units larger than about 100,000. In the quasi-experiment, we are unable to estimate the functional form of the cost curve. It may be that returns to scale are not linear but occur when the size exceeds certain thresholds. Another possibility is that the form of the cost curve depends on the scope of local policy tasks. For smaller portfolios (in the extreme, single-purpose units), economies of scale may be smaller, or disappear for large units. Finally, it should also be noted that cost curves can change over time. Diffusion of inventions (e.g., information technology) may change the relationship between size and costs. Such factors may influence the functional form of the cost curve.

We end by returning to the search for the optimal size for political systems. As indicated in the beginning of our article, this search may be illusory since there are dilemmas involved in designing government systems. Our findings show that the costs of running the system may be reduced by increasing jurisdiction size. But increasing jurisdiction size may also have costs. Welfare gains may be sacrificed because reducing the number of municipalities makes it more difficult to tailor public services to local preferences and offers citizens fewer tax-service packages to choose from when they decide where to live. Democracy may also suffer. Another analysis of the Danish quasi-experiment shows that citizens' internal political efficacy has dropped by nearly two points on a 10-point scale as a result of the municipal amalgamations (Lassen and Serritzlew 2011). Is a 10% reduction of the costs of running the system worth a considerable loss in internal political efficacy? This is a political question, and no scientist, not even a political one, can provide an objective answer to it. But scientists can inform politicians about the trade-offs involved in the architecture of government.

[9]We thank one of the referees for pointing this out to us.

Appendix

TABLE A1 Means for Dependent Variables and Control Variables for Amalgamated and Nonamalgamated Municipalities by Year

	2005	2006	2007	2008	2009	2010	2011
Administrative Costs per Capita							
Not amalgamated	5782	5833	5893	6067	6273	6255	6064
Amalgamated	5254	5354	5427	5403	5357	5259	5135
All municipalities	5411	5497	5566	5601	5630	5555	5412
Dispersal of Settlement							
Not amalgamated	1.81	1.81	1.81	1.81	1.81	1.89	1.89
Amalgamated	2.69	2.69	2.69	2.69	2.69	2.71	2.71
All municipalities	2.43	2.43	2.43	2.43	2.43	2.46	2.46
Social Housing							
Not amalgamated	25.5	25.5	25.5	23.4	23.4	24.3	24.3
Amalgamated	21.0	21.0	21.0	18.8	18.8	20.1	20.1
All municipalities	22.3	22.3	22.3	20.2	20.2	21.4	21.4
Children in One-Parent Families							
Not amalgamated	18.0	18.3	18.5	18.5	18.9	19.2	19.1
Amalgamated	15.1	15.4	15.6	15.5	15.9	16.4	16.7
All municipalities	16.0	16.3	16.5	16.4	16.8	17.2	17.4
Fiscal Pressure							
Not amalgamated	93.8	95.1	96.4	96.4	95.5	95.5	97.4
Amalgamated	102.6	102.3	102.7	103.0	103.1	102.8	102.0
All municipalities	100.0	100.2	100.8	101.1	100.8	100.6	100.6
Decentralization of Administrative Staff							
Not amalgamated	43.6	47.5	46.2	46.2	47.6	46.4	45.9
Amalgamated	41.3	46.5	45.1	45.1	47.2	45.8	44.9
All municipalities	42.0	46.8	45.4	45.4	47.3	46.0	45.2

Note: N each year—amalgamated: 66; not amalgamated: 28.

References

Alesina, Alberto, and Enrico Spolaore. 2005. *The Size of Nations.* Cambridge, MA: MIT Press.

Andrews, Rhys, and George A. Boyne. 2009. "Size, Structure and Administrative Overheads: An Empirical Analysis of English Local Authorities." *Urban Studies* 46(4): 739–59.

Baldersheim, Harald, and Lawrence E. Rose. 2010. "Territorial Choice: Rescaling Governance in European States." In *Territorial Choice: The Politics of Boundaries and Borders,* ed. Harald Baldersheim and Lawrence E. Rose. Houndmills, UK: Palgrave, 1–21.

Beck, Nathaniel, and Jonathan N. Katz. 1995. "What to Do (and Not to Do) with Time-Series Cross-Section Data." *American Political Science Review* 89(3): 634–47.

Benton, J. Edwin, and Darwin Gamble. 1984. "City/County Consolidation and Economies of Scale: Evidence from a Time-Series Analysis in Jacksonville, Florida." *Social Science Quarterly* 65(1): 190–99.

Besley, Timothy, and Anne Case. 2000. "Unnatural Experiments? Estimating the Incidence of Endogenous Policies." *Economic Journal* 110(467): F672–94.

Bhatti, Yosef, and Kasper Møller Hansen. 2011. "Who Marries Whom? The Influence of Societal Connectedness, Economic and Political Homogeneity, and Population Size on Jurisdictional Consolidations." *European Journal of Political Research* 50(2): 212–38.

Bhatti, Yosef, Asmus Leth Olsen, and Lene Holm Pedersen. 2011. "Keeping the Lights On: Citizen Service Centers in Municipal Amalgamations." *Administration in Social Work* 35(1): 3–19.

Bish, Robert L. 2001. *Local Government Amalgamations: Discredited Nineteenth-Century Ideals Alive in the Twenty-First.* C. D. Howe Institute Commentary No. 150. Toronto, Canada: C. D. Howe Institute.

Blom-Hansen, Jens. 2004. "Stordriftsfordele i den kommunale serviceproduktion? Skoleområdet som eksempel." In *Politisk Ukorrekt,* ed. Jens Blom-Hansen, Asbjørn Sonne Nørgaard, and Thomas Pallesen. Aarhus, Denmark: Aarhus Universitetsforlag, 260–74.

Boadway, Robin, and Anwar Shah. 2009. *Fiscal Federalism.* Cambridge: Cambridge University Press.

Boyne, George A. 1995. "Population Size and Economies of Scale in Local Government." *Policy and Politics* 23(3): 213–22.

Byrnes, Joel, and Brian Dollery. 2002. "Do Economies of Scale Exist in Australian Local Government? A Review of Research Evidence." *Urban Policy and Research* 20(4): 391–414.

Caulfield, Janice, and Helge O. Larsen. 2002. "Introduction." In *Local Government at the Millenium,* ed. Janice Caulfield and Helge O. Larsen. Opladen, Germany: Leske and Budrich, 9–27.

Christoffersen, Henrik, and Martin Paldam. 2003. "Markets and Municipalities: A Study of the Behavior of the Danish Municipalities." *Public Choice* 114(1/2): 79–102.

Cook, Thomas D., and Donald T. Campbell. 1979. *Quasi-Experimentation: Design & Analysis Issues for Field Settings.* Boston: Houghton Mifflin.

Dahl, Robert A. 1967. "The City in the Future of Democracy." *American Political Science Review* 41(4): 953–70.

Dahl, Robert A., and Edward R. Tufte. 1973. *Size and Democracy.* Stanford, CA: Stanford University Press.

Deller, Steven C., David L. Chicoine, and Norman Walzer. 1988. "Economies of Size and Scope in Rural Low-Volume Roads." *Review of Economics and Statistics* 70(3): 459–65.

Derksen, Wim. 1988. "Municipal Amalgamation and the Doubtful Relation between Size and Performance." *Local Government Studies* 14(6): 13–47.

Dollery, Brian, and Euan Fleming. 2006. "A Conceptual Note on Scale Economies, Size Economies and Scope Economies in Australian Local Government." *Urban Policy and Research* 24(2): 271–82.

Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach.* Cambridge: Cambridge University Press.

Hirsch, Werner Z. 1959. "Expenditure Implications of Metropolitan Growth and Consolidation." *Review of Economics and Statistics* 41(3): 232–41.

Houlberg, Kurt. 2000. *Kommunale stordriftsfordele—Hvor finder vi dem, og hvor store er de?* Copenhagen: AKF Forlaget.

Houlberg, Kurt. 2011. "Administrative stordriftsfordele ved kommunalreformen i Danmark. Sandede eller tilsandede?" *Scandinavian Journal of Public Administration* 15(1): 41–61.

King, David N., and Yue Ma. 2000. "Local Authority Size in Theory and Practice." *Environment and Planning C: Government and Policy* 18(3): 255–70.

Kohler, Ulrich, and Frauke Kreuter. 2005. *Data Analysis Using Stata.* College Station, TX: Stata Press.

Lassen, David Dreyer, and Søren Serritzlew. 2011. "Jurisdiction Size and Local Democracy: Evidence on Internal Political Efficacy from Large-Scale Municipal Reform." *American Political Science Review* 105(2): 238–58.

Lewis, Paul G. 2011. "Size and Local Democracy: Scale Effects in City Politics." *PS: Political Science and Politics* 44(1): 107–9.

Martins, M. R. 1995. "Size of Municipalities, Efficiency, and Citizen Participation: A Cross-European Perspective." *Environment and Planning C: Government and Policy* 13(4): 441–58.

Meyer, Bruce D. 1995. "Natural and Quasi-Experiments in Economics." *Journal of Business and Economic Statistics* 13(2): 151–61.

Mouritzen, Poul Erik. 1999. *De små kommuner: De er ikk' så ringe endda?* Copenhagen: AKF Forlaget.

Mouritzen, Poul Erik, ed. 2006. *Stort er godt: Otte fortællinger om tilblivelsen af de nye kommuner.* Odense, Denmark: Syddansk Universitetsforlag.

Mouritzen, Poul Erik. 2010. "The Danish Revolution in Local Government: How and Why?" In *Territorial Choice: The Politics of Boundaries and Borders,* ed. Harald Baldersheim and Lawrence E. Rose. Houndmills, UK: Palgrave, 21–41.

Newton, K. 1982. "Is Small Really So Beautiful? Is Big Really So Ugly? Size, Effectiveness, and Democracy in Local Government." *Political Studies* 30(2): 190–206.

Oates, Wallace E. 1972. *Fiscal Federalism.* New York: Harcourt Brace Jovanovich.

Ostrom, Elinor. 1972. "Metropolitan Reform: Propositions Derived from Two Traditions." *Social Science Quarterly* 53(3): 474–93.

Regeringen and Dansk Folkeparti. 2004. *Aftale om strukturreform.* Copenhagen: Ministry of Interior and Health.

Sancton, Andrew. 2000. *Merger Mania: The Assault on Local Government.* Montreal, Canada: McGill-Queen's University Press.

Sawyer, Malcolm C. 1991. *The Economics of Industries and Firms: Theories, Evidence and Policy.* London: Routledge.

Scherer, Frederic M., and David Ross. 1990. *Industrial Market Structure and Economic Performance.* Boston: Houghton Mifflin.

Sekhon, Jasjeet S., and Roeio Titiunik. 2012. "When Natural Experiments Are Neither Natural nor Experiments." *American Political Science Review* 106(1): 35–57.

Sharpe, L. J. 1970. "Theories and Values of Local Government." *Political Studies* 18(2): 153–74.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditure." *Journal of Political Economy* 64(5): 416–24.

Tiebout, Charles M. 1960. "Economies of Scale and Metropolitan Governments." *Review of Economics and Statistics* 42(4): 442–44.

Treisman, Daniel. 2007. *The Architecture of Government: Rethinking Political Decentralization.* Cambridge: Cambridge University Press.

Vetter, Angelika, and Norbert Kersting. 2003a. "Democracy versus Efficiency? Comparing Local Government Reforms across Europe." In *Reforming Local Government in Europe,* ed. Norbert Kersting and Angelika Vetter. Opladen, Germany: Leske and Budrich, 11–29.

Vetter, Angelika, and Norbert Kersting. 2003b. "Reforming Local Government: Heading for Efficiency and Democracy." In *Reforming Local Government in Europe,* ed. Norbert Kersting and Angelika Vetter. Opladen, Germany: Leske and Budrich, 333–49.

Warren, Mark E. 2011. "Voting with Your Feet: Exit-Based Empowerment in Democratic Theory." *American Political Science Review* 105(4): 683–701.

Williams, Rick L. 2000. "A Note on Robust Variance Estimation for Cluster-Correlated Data." *Biometrics* 56(2): 645–46.

Wooldridge, Jeffrey M. 2009. *Introductory Econometrics.* London: South-Western Cengage Learning.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1: Measurement of variables
Table S2: Conversion of administrative costs to 2011 task portfolio
Table S3: Conversion of administrative staff to 2011 task portfolio