Economica (2003) **70**, 673–690

Size and School District Consolidation: Do Opposites Attract?

By DAVID M. BRASINGTON*

Final version received 15 August 2002.

Municipalities sometimes retain separate police departments and park services while cooperating in public schooling services with neighboring municipalities. The theoretical model of Ellingsen (*Journal of Public Economics*, **68**, 251–68, 1998) predicts that: (1) under Tiebout sorting, larger size differences make big municipalities more likely to consolidate with small ones, but small municipalities less likely to consolidate with big ones; (2) municipalities never excessively consolidate. The current study examines 298 pairs of municipalities that could consolidate schooling. The decision-making process of the larger and smaller member of each pair is examined separately. The Poirier bivariate probit results are consistent with Ellingsen's predictions but contradict previous empirical findings.

INTRODUCTION

In the 1937 school year the United States had 119,001 public school districts. By the 1997 school year there were 14,805.

Some of the decrease may stem from annexation, in which two independent cities form one city, sharing schooling services, police services, and political representation.[1] But much of the decrease may stem from consolidation. In school district consolidation, two independent cities may retain separate control over police services and political representation, yet share public schooling services. Consolidation of local public good provision is also called integration of services, contracting for services, jointly providing services, and cooperating in the provision of services.

Until recently economists largely ignored consolidation. One of the many issues that deserves more attention is the role of size in the consolidation decision. Consider a pair of cities that could potentially consolidate schooling. Call this a potential consolidation pair. One member of the potential consolidation pair is more populous than the other. Is the bigger city more likely to approve consolidation the smaller is its consolidation partner? Does a greater difference in size make the smaller city more likely to approve consolidation? And what is the end result: does a bigger difference in size promote or hinder consolidation?

The current study presents an empirical test of the role of size in the decision of cities to consolidate public schooling provision. The empirical test is inspired by the theoretical model of Ellingsen (1998). Ellingsen's model predicts that under plausible conditions large cities would want to consolidate with small cities, but small cities might not want to consolidate with big cities. If both cities must approve consolidation, increasing the difference in size between two cities makes consolidation less likely. Ellingsen's model also predicts that there will never be excessive consolidation: that cities will never consolidate when it leads to a welfare loss, but they may even forego a consolidation that would enhance total welfare. Using a data set from Ohio of

298 potential consolidation pairs, the current study is the first to examine how the larger and smaller members of each potential consolidation pair decide whether to consolidate with a city of different size.

The empirical test shows that the determinants of the consolidation decision for the larger and smaller member of each potential consolidation pair are similar. The biggest difference is the role of size. The empirical test supports the predictions of Ellingsen's theoretical model. Bigger cities prefer to consolidate with smaller neighbors, but smaller cities do not want to consolidate with bigger cities. Smaller cities may feel they will lose their ability to influence school policy decisions by consolidating with a more populous neighbor. Therefore, all else constant, only cities of similar size tend to consolidate. Only two previous empirical studies examine the role of size in the consolidation decision; the results of the current study contradict those studies. The difference probably stems from the new way the estimations are performed, which allows a richer examination of the role of size.

The results suggest that cities with a larger tax base are more likely to retain sole control over schooling provision. Differences in income and racial composition also hinder consolidation. Finally, the paper provides evidence consistent with the second prediction of Ellingsen's model: there will never be excessive consolidation, but cities may forego a consolidation that would enhance total welfare.

The paper concludes by relating the school district consolidation decision to many of the decisions facing nations. The approach used in the current study might readily be applied to examining why nations reunite, split up, sign treaties, and join or withdraw from international organizations.

I. Theoretical Discussion

Some older work (Guesnerie and Oddou, 1981) and a flurry of recent theoretical work investigates the consolidation and fragmentation of public service provision (e.g., Austin, 1998; Brasington, 2003; Conley and Dix, 1999; Borck, 1998; Henderson and Thisse, 1997; Pecorino, 1999). These models typically discuss a tradeoff between the scale economy gains from consolidation on the one hand and the loss of sole control over education provision on the other hand. Despite the recent papers in this area, only the theoretical model of Ellingsen (1998) deals directly with the issue of primary concern for the current study: how size influences preferences for consolidation. A summary of the most relevant portions of his model are presented to motivate the empirical testing. The summary also points out the ways in which public schooling, the focus of the empirical section, fits with and deviates from Ellingsen's model.

In Ellingsen's model two cities choose to provide a public good independently (separation) or to cooperate in provision with the other city (consolidation). The separation/consolidation decision is decided by a majority vote in each city. This is essentially a median voter model because the median voter is decisive in the consolidation decision, and consolidation may change the identity of the median voter in the new school district. Each city has veto power; both cities must agree to consolidate. People have rational expectations, so they foresee the outcomes of both separation and consolidation. In the first period people vote on consolidation, and in the second period they make policy decisions conditional on whether provision is separate or consolidated.

Metropolitan areas in the U.S. generally consist of a large central city with a number of suburban cities surrounding it. Tiebout (1956) says people move to the city within a metropolitan area that most closely matches their preferences for public goods. Ellingsen's two-city model follows that of Tiebout in that people with a stronger preference for the public good all live in one city, while those with a weaker preference for the public good live in the other city. The cities are allowed to differ in size. Therefore the level of public good provision can differ, reflecting taste and size differences.

Most of Ellingsen's model focuses on locally-provided pure public goods, but his local public goods extension is more relevant for schooling. For people of type j and preferences θ, the following utility function describes the local public goods extension for the provision of private good x and public good z in cities $i = \{X, Y\}$:

$$(1) \qquad U_j = x_j + \theta_j v(\lambda^i z^X + \gamma^i z^Y)$$

In Equation (1), λ and γ allow the benefits of z provision to vary across production sites. λ^X measures the benefits X gets from production in X, and γ^X is the spillover benefit city X gets from provision z^Y in city Y. Similarly, γ^Y is the benefit to city Y of its own provision z^Y, while λ^Y measures the spillover benefit to city Y of provision z^X in city X.

Public schooling may be associated with positive spillovers, the case in Equation (1) where $\lambda^Y > 0$ and $\gamma^X > 0$. For instance, Brasington (2000) models education as a joint product, with the pure private component consumed internally but the pure public component of schooling spilling over into neighboring communities. His empirical test suggests that the levels of crime and school amenities in one district affect a neighboring school district's proficiency test outcomes. Wyckoff (1984) also finds evidence of a public component of education, and Haveman and Wolfe (1984) suggest that the external benefits to education may double its social rate of return. The provision of education may benefit non-recipients through greater workforce productivity; a more educated electorate; and improved community social capital, which facilitates community action, like collectively preventing a landfill from entering (Stiglitz, 2000; Fischel, 2002).[2]

Ellingsen (1998) shows that consolidation often improves welfare, but the incentive to free ride may prevent it. Restricting attention to the Tiebout sorting cases in which the public good has a positive, rather than a negative externality, Ellingsen's model has the predictions shown in Table 1.

The predictions of the model are the same for Case 1 in which tastes are identical and Case 2 in which the big city provides more public good than the small. The predictions differ in Case 3 in which the big city provides less public good than the small. But in the case of schooling (the focus of the empirical section), Case 3 is unlikely. On a per-capita basis a small city might provide more schooling than a big city if the taste differences are big enough, but looking at aggregate amounts of schooling provided, a bigger city usually

TABLE 1
PREDICTIONS OF ELLINGSEN'S (1998) MODEL

Case	Tastes	Excessive integration?	Big wants consolidation?	Role of size differences
1	Identical	No	Always	↑size differences → ↓probability of consolidation
2	Big city provides more than small	No	Always	↑size differences → ↓probability of consolidation
3	Big city provides less than small	No	Not always	↑size differences → ↑probability of consolidation

provides more. In fact, for the sample at hand, an investigation suggests that the larger city provides a larger amount of public schooling than the smaller city over 95% of the time.[3] Case 3 does not describe the data well. Attention is restricted to Case 1 and Case 2.

Cases 1 and 2 predict that there will not be excessive integration. That is, consolidation will not take place when it detracts from welfare, but even when consolidation is welfare-enhancing cities will sometimes choose not to consolidate. The model also predicts that the bigger city wants to consolidate with a smaller city, but that a smaller city will not always want to consolidate with a bigger city. The big city tends to provide more public good than the small is willing to pay for, or less than the small is willing to accept. In fact, all else constant, increasing the size difference between the big and small city makes the small city more averse to consolidation. When the small city has veto power, a larger size difference makes consolidation less likely.

The predictions of Ellingsen's model hold for both pure public goods and local public goods with spillovers. The smaller are the spillover terms λ^Y and γ^X in Equation (1), the less likely it is that one community will completely free ride on the production of its neighbor. In fact, even when the communities consolidate, having smaller spillovers makes it more likely that both communities will provide z. Such is often the case for public schooling. Members of a consolidated school district generally share a high school and a school board, but have elementary schools in each community.

Ellingsen's model assumes there are no congestion costs, and his results may rely on this assumption. It is clear, however, that adding pupils will raise a school's costs, so the current study cannot serve as a direct test of Ellingsen's model. However, Ellingsen's model motivates some important questions. Do larger cities prefer consolidating with smaller cities? Do smaller cities dislike consolidating with bigger cities? And how does increasing the *difference* in size between two cities affect their willingness to consolidate? Is there any evidence of excessive integration? These are questions that can be explored by empirical analysis.

II. PREVIOUS EMPIRICAL WORK

It seems only two empirical studies address the issue of size and consolidation: Ferris and Graddy (1988) and Brasington (1999b). Even so, neither study directly addresses how larger differences in size affect the consolidation decision of the larger and smaller city.

Ferris and Graddy (1988) investigate what makes counties and cities contract for service provision. The contracting partner can be private firms, non-profit organizations, or other local governments. Ferris and Graddy focus on eight local services: street lights, street repair, public health, hospitals, bus services, programs for the elderly, recreational services, and residential waste collection. For each of these public services Ferris and Graddy perform an ordered multinomial logit with the following dependent variables: the service is produced internally, the service is contracted out entirely, and the service is produced both internally and externally (joint provision).

Ferris and Graddy's explanatory variables include population and its square. For public health and programs for the elderly they find population negative and its square positive. Ferris and Graddy interpret this to mean local governments with large and small populations are most likely to provide these services jointly and to contract externally, with medium-sized local governments most likely to retain local control. They reason that smaller local governments contract for services because they have the most to gain from scale economies (Andrews et al., 2002; Duncombe and Yinger, 2000). The largest local governments contract for services because they have the most to gain from competition among service suppliers. However, Ferris and Graddy do not investigate how increasing the *difference* in size between two potential contracting partners affects the willingness of big and small cities to contract.

Brasington (1999b) investigates the determinants of cooperation in the provision of public schooling, the public service that commands the most money, and one of the few public services not investigated in Ferris and Graddy (1988). Brasington includes *DIFFERENCE IN PUPILS* to capture the difference in size between two neighboring school districts. *DIFFERENCE IN PUPILS* is positive in his regression, and Brasington interprets this to mean that consolidation is more likely between local governments of different size.

Brasington's results show a general tendency for large and small cities to consolidate, but his estimation cannot disentangle how large and small cities make their separate consolidation decisions. As such, it cannot say how large and small cities feel about consolidating with a city of different size. The problem is discussed in greater detail later, but in short, Brasington arranges his data in a random fashion and constrains his parameter estimates, providing a single parameter estimate for two cities. Without the constraint, there would be a separate parameter estimate for city X and for city Y. Arranging cities X and Y by size and allowing unconstrained parameter estimates is the only way to determine how differences in size affect the separate consolidation decisions of large and small cities. In fact, the results of the current study contrast with those of Brasington (1999b).

III. Institutional Background

The empirical analysis spans the six major metropolitan areas in the state of Ohio: Akron, Cincinnati, Cleveland, Columbus, Dayton, and Toledo.

Restricting the analysis to one state avoids the problem of controlling for differences in home rule laws and consolidation procedures across states (Brasington, 1999b).

Once a new local political jurisdiction like a city, village or township forms, the state constitution of Ohio mandates that it provide public education. It may form its own school district, continue to use the school district it used before, or it may consolidate with a different school district.[4]

Ohio law makes forming and disbanding a consolidated school district fairly easy. To consolidate, both communities' boards of education basically just need to vote in favor of consolidation; consolidation fails if one of the boards votes against it (Baldwin's Ohio Revised Code, 1995).[5] After consolidation voters elect a new unified school board and the pre-consolidation school boards disband. If two communities cooperate in the provision of schooling, it must be for all grade levels: kindergarten through twelfth grade. The communities must be contiguous; the law forbids consolidation between communities that do not share a geographical boundary.

Consolidation has gained and lost favor in waves. Between the 1930s and the 2002–2003 school year the number of school districts in Ohio fell from 1,936 to 612. Most of these mergers happened in the 1930s and the 1960s. In the 1930s wave, county offices encouraged consolidation, while the State Board of Education encouraged consolidation in the 1960s wave. In neither episode were legal threats or financial incentives used to promote consolidation. Since 1985 three consolidated school districts formed, one school district disbanded for lack of students, and two split into their pre-consolidation components (Brasington, 2003).

The analysis excludes central cities for two reasons. First, their boundaries are usually historical artifacts rather than conscious choice based on Tiebout sorting. Inflexibilities may prevent the redrawing of jurisdictional boundaries (Hoyt, 1999; Garasky and Haurin, 1997). Second, the theoretical model examines the consolidation decision between communities shaped by Tiebout sorting, so communities are relatively internally homogeneous. Central cities have much more demographic diversity than the suburban communities that surround them.

Before 1955, if territory was annexed, residents had a new mayor and were automatically assigned to the school district of the city that annexed them. But in 1955 the law changed so that although municipal boundaries changed, the school district assignment did not necessarily change. Consequently, political jurisdiction boundaries do not always coincide with school district boundaries. Data constraints limit the analysis to jurisdictions that mainly send their children to a single school district; any jurisdiction whose geographical school assignment is markedly split among more than one school district is excluded from the sample.[6]

To restrict the sample to urban communities, any school district on the edge of the metropolitan area that includes a large portion of rural land is omitted.[7] Consolidated school districts typically have schools in each of their member communities. Sometimes very small communities have no school buildings and are therefore said to contract out for educational services. These small communities are omitted if a larger community completely encompasses them; however, they are kept if the small community can choose between contracting partners. By looking at maps for metropolitan communities that either clearly belong to a consolidated school district or clearly provide schooling independently, 298 potential consolidation pairs are found. Eighty-three (28%) of the potential consolidation pairs form consolidated school districts.

Local property taxes provide slightly over half of Ohio school districts' revenues. State government grants contribute roughly one-third of revenues, and the rest comes from miscellaneous sources, the national government, and school district income taxes. Approximately 123 of Ohio's 612 school districts have school district income taxes. Communities with school district income taxes tend to be rural. In fact, only one school district in the sample had a school district income tax in the year of the sample.

IV. Data

The data is a cross-sectional look at which municipalities provide schooling independently or as part of a consolidated school district in 1990. Even if the decision to consolidate occurred in 1967, the consolidation status of a municipality in 1990 should reflect its preferences fairly accurately, because Ohio law makes it easy to consolidate and to undo consolidation. The decision to consolidate depends on size, property value, and socio-economic factors. Many studies find scale economies in public schooling, both at the school level and the school district level.[8] In fact, the most frequently-cited reason to consolidate schooling is to gain scale economies (Kenny and Schmidt, 1994). To allow for economies of scale in schooling provision, *#PUPILS* is included in the analysis; it represents the number of school-aged children in the community. *#PUPILS SQUARED* is also included. The smallest cities may effectively be forced to consolidate to gain adequate scale economies in schooling provision, but as *#PUPILS* increases, a city may choose separation to maintain political control over schooling. *#PUPILS* is therefore expected to be negatively related to consolidation. However, as a city grows increasingly large, it becomes more likely to retain control in a consolidated district, and it might still reap some scale economies. *#PUPILS SQUARED* is therefore expected to be positively related to consolidation.

Economies of scale are also captured by *POTENTIAL PUPILS*; it represents the number of school-aged children there would be if the two communities consolidated their school districts. As long as there are scale economies to reap, *POTENTIAL PUPILS* is expected to raise the probability of consolidation.[9]

PROPERTY VALUE is the value of real estate in the community. High levels of *PROPERTY VALUE* may discourage consolidation. The larger the community's tax base, the more easily the community's own most-desired schooling level can be funded. If each community can easily fund its own schools, consolidating with a neighbor will not help the community financially, but it will force schooling provision levels away from the individual levels most preferred before consolidation (e.g., Borck, 1998).

The tax price depends on the tax rate and the tax base. If one community has more per-pupil property value than its neighbor, consolidation would dilute the tax base and raise the tax price. The community with less property value might like to consolidate with its property-rich neighbor, but the feeling is probably not mutual. *DIFFERENCE IN PROPERTY VALUE* is therefore expected to reduce the likelihood of consolidation.

Ellingsen (1998) describes different types of residents having different preferences over public good provision. Resident types may be reflected by racial composition, incomes, and education levels, all of which may affect preferences for schooling. Therefore school district consolidation may depend on socioeconomic differences like *DIFFERENCE IN %WHITE*, *DIFFERENCE IN INCOME*, and *DIFFERENCE IN EDUCATION*. The larger the socioeconomic differences, the less likely a merger.[10] Because the median voter is decisive in Ellingsen's model, the *DIFFERENCE IN INCOME* variable is calculated using median income in each municipality.

Central to the model is the tradeoff between scale economies and preferences, and preferences may be expressed by the level of expenditures chosen. But preferred expenditure levels are measured at the school district level, and are not observable at the municipality level. Still, desired expenditures are driven by a city's size, property value, and socio-demographic characteristics, which are included in the empirical model.

The focus of Ellingsen's theoretical model is the relationship between differences in size and consolidation. Ferris and Graddy (1988) and Brasington (1999b) provide valuable insights on how cooperation in public service provision between communities is affected by size differences. The current study uses a different approach than prior studies. It looks at pairs of communities that could consolidate, called potential consolidation pairs. One of the two communities is larger, the other is smaller. The current study examines what affects the larger community's consolidation decision and the smaller community's decision. By splitting the sample into larger and smaller communities the current study provides the first direct look at the influence of size differences on the separate consolidation decisions of smaller and larger communities.

Let *DIFFERENCE IN PUPILS* be the difference in the number of pupils in each community in a potential consolidation pair. The theoretical model of Ellingsen suggests that *DIFFERENCE IN PUPILS* should encourage the larger community in each pair to consolidate, but make the smaller community in each pair resist consolidation. The variables used, along with their definitions and sources, are shown in Table 2.

V. Empirical Estimation Technique

The following equation is estimated:

$$\text{CONSOLIDATED} = f(S_i, V_i, D_i) \qquad (2)$$

where S represents size variables for each member i of a potential consolidation pair, V represents property value, and D represents demographic differences. CONSOLIDATED is a dummy variable that takes the value 1 if two adjacent communities both decide to form a consolidated school district; it takes the value 0 if at least one of the communities chooses separation.

TABLE 2
VARIABLE DEFINITIONS AND SOURCES

Variable	Definition	Source
#PUPILS	Number of school-aged children in the municipality, in hundreds of thousands	1
DIFFERENCE IN PUPILS	*#PUPILS* in the community in question minus *#PUPILS* in potential consolidation partner; calculated in absolute value	1
POTENTIAL PUPILS	*#PUPILS* in the community in question plus *#PUPILS* in potential consolidation partner	1
PROPERTY VALUE	Assessed valuation of real estate and public utility values from 1989 abstracts and 1990 collections, plus personal tangible value from 1990 collections, in billions of dollars	2
DIFFERENCE IN PROPERTY VALUE	*PROPERTY VALUE/#PUPILS* in municipality in question minus *PROPERTY VALUE/#PUPILS* of potential consolidation partner; calculated in absolute value	1,2
DIFFERENCE IN INCOME	Median household income, in hundreds of thousands of dollars, for the community in question minus that of its potential consolidation partner; calculated in absolute value	1
DIFFERENCE IN %WHITE	Proportion of white residents in the community in question minus the proportion of white residents in the neighboring jurisdiction; calculated in absolute value	1
DIFFERENCE IN EDUCATION	Proportion of residents who have at least attended college in the community in question minus the proportion of residents who have at least attended college in the neighbouring jurisdiction; calculated in absolute value	1
CONSOLIDATED	Dummy variable for whether the two neighbouring jurisdictions cooperate in public schooling provision; that is, whether they form a consolidated school district	3

Sources: 1 = US Bureau of the Census (1990); 2 = Ohio Municipal Advisory Council (1993); 3 = Ohio Department of Education (1985).

Estimation of Equation (2) requires a statistical technique that allows for partial observability: only the product of the two neighboring communities' consolidation choices is observed, not the individual choices. Consolidation is observed only when both neighbors choose to do it. Separation is observed if community X, community Y, or both communities choose not to consolidate. In addition, the consolidation decisions of the communities are simultaneously determined.

The statistical technique that matches the current situation is the Poirier bivariate probit (Poirier, 1980). It uses full information maximum likelihood to allow for correlation between the disturbance terms of the communities' decisions. Poirier's partial observability model has the following log-likelihood:

$$(3) \qquad \ln L = \sum\nolimits_{z=1} \ln \Phi[\beta_1' \boldsymbol{x}_{i1}, \beta_2' \boldsymbol{x}_{i2}, \rho] + \sum\nolimits_{z=0} \ln(1 - \Phi[\beta_1' \boldsymbol{x}_{i1}, \beta_2' \boldsymbol{x}_{i2}, \rho])$$

where z is each neighboring community's final decision on whether to consolidate, Φ is the cumulative distribution function, assumed to be bivariate standard normal; the $\boldsymbol{x}$'s $= \{S,V,D\}$ are the factors underlying the decision, and ρ is the correlation between the two communities' decisions.[11]

The log-likelihood function of the Poirier bivariate probit is not globally convex. Therefore the negative of the log-likelihood is minimized using the Davidon–Fletcher–Powell algorithm. This algorithm does not use second derivatives, and it has excellent convergence properties, even for ill-behaved problems (Greene, 1997, p. 204).

The arrangement of the data is fundamental to the way the current study contributes to the literature. To better understand how the data are arranged, consider Figure 1. The illustration shows four political jurisdictions in a metropolitan area. Each jurisdiction can maintain its own school district or it can consolidate with a willing neighbor. Table 3 shows how the data are arranged for *DIFFERENCE IN PUPILS* for each potential consolidation pair.

In Figure 1 only A and D consolidate. A and C are not adjacent. They cannot legally form a consolidated school district with each other; therefore, there is no potential consolidation pair observation between A and C directly. However, because A and D consolidate their schools, C may try to join the consolidated school district A&D. C may merge with A, then, but only through the actual consolidated school district consisting of A and D. In sum, observations are only permitted between (1) adjacent municipalities, and (2) municipalities and adjacent consolidated school districts. Such a rule reduces the excessive number of possible combinations that might otherwise arise, such as C&B merging with A&D.

Each jurisdiction appears in multiple observations, but the observations are assumed to be independent. The assumption of independence is untenable: A's decision to merge with D affects A's decision to merge with B. The data setup creates a cluster of non-zero error covariances for each jurisdiction, causing inconsistent parameter estimates and biased standard errors. The severity of the problem is unknown. However, if each jurisdiction appears in three or four observations, this is a small number relative to the 298 observations in the sample, so one may hope that the bias does not affect the results much.

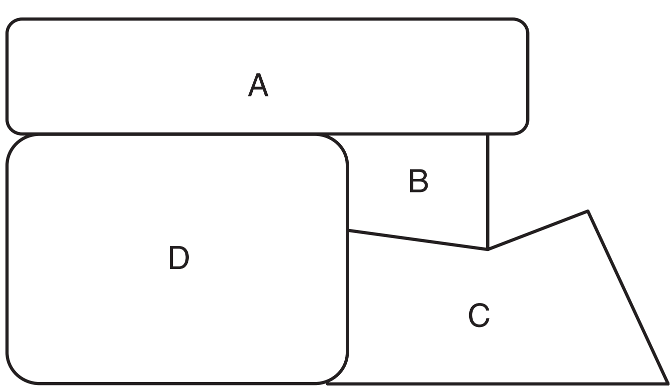

FIG. 1.

TABLE 3

Potential consol. pair	Entity 1	Entity 2	Merged	*#PUPILS 1*	*#PUPILS 2*	[Difference Pupils 1]	[Difference Pupils 2]
1	A	B	0	3500	2000	1500	1500
2	D	A	1	4750	3500	1250	1250
3	D	B	0	4750	2000	2750	2750
4	D	C	0	4750	4500	250	250
5	C	B	0	4500	2000	2500	2500
6	C	A&D	0	4500	8250	3750	3750

TABLE 4
MEANS[a]

Variable	Bigger city	Smaller city
#PUPILS	0.045	0.017
	(0.037)	(0.016)
DIFFERENCE IN PUPILS	0.028	0.028
	(0.032)	(0.032)
POTENTIAL PUPILS	0.062	0.062
	(0.048)	(0.048)
PROPERTY VALUE	0.46	0.22
	(0.34)	(0.19)
DIFFERENCE IN PROPERTY VALUE	0.056	0.056
	(0.20)	(0.20)
INCOME	0.43	0.43
	(0.20)	(0.19)
DIFFERENCE IN INCOME	0.0028	0.0028
	(0.22)	(0.22)
%WHITE	0.87	0.87
	(0.18)	(0.22)
DIFFERENCE IN %WHITE	0.0017	0.0017
	(0.21)	(0.21)
EDUCATION	0.53	0.52
	(0.16)	(0.17)
DIFFERENCE IN EDUCATION	0.0115	0.0115
	(0.16)	(0.16)

[a]Means shown with standard deviation in parentheses below for the more populous and less populous municipality in each pair of municipalities that could potentially consolidate. Number of pairs = 298. Mean of *CONSOLIDATED* is 0.28.

With the data arranged as shown the Poirier bivariate probit may proceed as follows:

1. Dependent variable = CONSOLIDATED
2. Right-hand side 1: DIFFERENCE IN PUPILS 1, PROPERTY VALUE 1....
3. Right-hand side 2: DIFFERENCE IN PUPILS 2, PROPERTY VALUE 2....

Note the way the chart arranges the data. The first community in the potential consolidation pair is always bigger, so 'right-hand side 1' is the big city and 'right-hand side 2' is the small city. Arranging the data by size shows how the bigger and smaller communities in each potential consolidation pair form their separate consolidation decisions. In contrast, Brasington (1999b) arranges his data randomly. The random arrangement allows Brasington to restrict the parameter estimates of *DIFFERENCE IN PUPILS* to be equal, simplifying the exposition. But only by arranging the data by community size and running an unrestricted regression can a researcher determine how size differences affect big and small communities' consolidation decisions and search for other differences in how larger and smaller communities make their individual consolidation decisions.

Table 4 shows variable means for the larger and smaller community in each potential consolidation pair.

VI. Empirical Results

The results of the Poirier bivariate probits appear in Table 5.

The "Size Only" columns show results when only variables related to the size of a city are included. The results for both large and small cities are similar. #PUPILS is negative, and its square is positive. As a city gets larger the economies of scale gains of consolidation tempt it less; the city instead prefers to enjoy complete political control over the provision of public schooling. But beyond a certain size consolidation becomes attractive again. A big city can still gain some scale economies without sacrificing political control. This is consistent with the prediction of Ellingsen's (1998) theory that big cities always favor consolidating with smaller cities.[12] *POTENTIAL PUPILS* is positive: the lure of consolidation grows the more scale economies there are to gain.

DIFFERENCE IN PUPILS is the variable of focus. For the bigger city *DIFFERENCE IN PUPILS* is positive. The larger the population gap between the big city and the small city, the more the big city wants to consolidate. The big city gains scale economies without sacrificing political control over schooling decisions.

But for the smaller city the parameter estimate of *DIFFERENCE IN PUPILS* is negative. All else constant size differences discourage small cities from merging with big cities; the bigger are the big cities, the more the small cities resist consolidation. This result is also consistent with the predictions of Ellingsen's (1998) theory. However, it differs from Brasington (1999b). Brasington's constraining of the parameter estimates and random sorting of the data are responsible for the difference. Only by sorting the data by size and running an unconstrained regression for both the big and small members of a consolidation pair can the role of size differences in the consolidation decision be examined in detail.

The "Size Only" columns of results are suggestive, but will they hold up when additional controls are added? The "Add Property Value" columns of results in Table 5 say they do. *DIFFERENCE IN PUPILS* is still positive for big cities and negative for small cities. All else constant, a big city is encouraged to consolidate the smaller the small city is. Decreasing the size gap between the

TABLE 5
RESULTS[a]

Variables	Size only		Add property value		Add all demographics		Add income only	
	Big city	Small city	Big city	Small city	Big city	Small city	Big city	Small city
#PUPILS	− 44.9**	− 47.0**	− 41.2**	− 39.9**	− 41.0**	− 40.8**	− 41.0**	− 41.8**
	(18.0)	(14.6)	(15.3)	(10.5)	(10.9)	(7.7)	(15.5)	(11.2)
#PUPILS SQUARED	47.6**	125.1**	41.2**	97.5*	43.1**	122.3*	43.5**	119.2**
	(4.2)	(2.7)	(3.6)	(2.0)	(2.7)	(1.8)	(3.8)	(2.5)
DIFFERENCE IN PUPILS	19.1**	− 19.6**	19.1**	− 19.6**	18.6**	− 18.8**	18.6**	− 19.3**
	(9.9)	(15.5)	(9.9)	(15.7)	(7.0)	(10.9)	(9.9)	(15.8)
POTENTIAL PUPILS	1.73**	1.71**	1.74**	1.70**	1.62**	1.59**	1.62**	1.63**
	(16.6)	(16.3)	(16.8)	(16.3)	(11.2)	(10.9)	(16.0)	(16.0)
PROPERTY VALUE	—	—	− 0.35**	− 0.54**	− 0.27*	− 0.45*	− 0.23**	− 0.43**
	—	—	(3.7)	(3.7)	(2.0)	(2.3)	(2.5)	(3.0)
DIFFERENCE IN PROPERTY VALUE	—	—	-0.8×10^{-3}	0.07	− 0.02	0.04	− 0.003	0.06
	—	—	(0.0)	(1.3)	(0.3)	(0.5)	(0.1)	(1.0)
DIFFERENCE IN INCOME	—	—	—	—	− 0.44**	− 0.44**	− 0.45**	− 0.46**
	—	—	—	—	(3.8)	(3.8)	(6.1)	(6.3)
DIFFERENCE IN %WHITE	—	—	—	—	− 0.30**	− 0.31**	—	—
	—	—	—	—	(2.8)	(2.8)	—	—
DIFFERENCE IN EDUCATION	—	—	—	—	0.08	0.02	—	—
	—	—	—	—	(0.4)	(0.1)	—	—
CONSTANT	0.39**	0.36**	0.41**	0.37**	0.53**	0.51**	0.50**	0.48**
	(15.8)	(15.5)	(14.6)	(13.7)	(11.6)	(11.2)	(16.0)	(15.3)
Rho	− 0.76		− 0.98**		0.99		− 0.83*	
	(0.8)		(12.2)		(0.6)		(2.2)	
Pseudo *R*-square	1.41		1.40		1.90		1.41	
% separation correctly predicted	0.97		0.95		0.93		0.96	
% consolidations correctly predicted	0.44		0.45		0.97		0.45	

[a]Parameter estimates shown with standard error in parentheses below.
* = statistically significant at 0.10;
** = statistically significant at 0.01. Number of observations = 298. Pseudo *R*-square is that of McIntosh and Dorfman (1992). % separation correctly predicted is the proportion of times the regression correctly predicted that the two cities would remain independent; % consolidations correctly predicted is the proportion of times the regression correctly predicted that the two cities would consolidate. Dependent variable is *CONSOLIDATED*.

TABLE 6
ADDITIONAL RESULTS[a]

Variables	Add race only		Add education only		Add race and education		Add income and education	
	Big city	Small city	Big city	Small city	Big city	Small city	Big city	Small city
#PUPILS	– 40.0**	– 38.8**	– 40.8**	– 39.1**	– 39.6**	– 38.5**	– 42.3**	– 41.3**
	(10.4)	(7.3)	(15.3)	(10.4)	(14.8)	(10.3)	(16.0)	(11.0)
#PUPILS SQUARED	37.1*	101.9	38.2**	89.7*	36.0**	95.8*	45.2**	113.8**
	(2.3)	(1.5)	(3.3)	(1.9)	(3.2)	(2.0)	(4.0)	(2.4)
DIFFERENCE IN PUPILS	18.7**	– 19.0**	18.9**	– 19.5**	18.6**	– 19.0**	18.9**	– 19.3**
	(6.9)	(10.8)	(9.8)	(15.7)	(9.8)	(15.4)	(10.0)	(15.8)
POTENTIAL PUPILS	1.68**	1.64**	1.72**	1.69**	1.68**	1.65**	1.67**	1.64**
	(11.5)	(11.1)	(16.8)	(16.3)	(16.4)	(15.9)	(16.4)	(16.0)
PROPERTY VALUE	– 0.38**	– 0.55**	– 0.31**	– 0.52**	– 0.35**	– 0.54**	– 0.23**	– 0.44**
	(2.9)	(2.7)	(3.3)	(3.6)	(3.8)	(3.8)	(2.4)	(3.1)
DIFFERENCE IN PROPERTY VALUE	– 0.02	0.05	– 0.003	0.07	– 0.02	0.06	– 0.003	0.06
	(0.3)	(0.7)	(0.1)	(1.2)	(0.4)	(1.0)	(0.1)	(1.0)
DIFFERENCE IN INCOME	—	—	—	—	—	—	– 0.42**	– 0.41**
	—	—	—	—	—	—	(5.1)	(5.0)
DIFFERENCE IN %WHITE	– 0.33**	– 0.34**	—	—	– 0.28**	– 0.27**	—	—
	(3.2)	(3.3)	—	—	(3.6)	(3.5)	—	—
DIFFERENCE IN EDUCATION	—	—	– 0.41**	– 0.48**	– 0.24*	– 0.32**	– 0.12	– 0.19
	—	—	(3.5)	(4.0)	(1.9)	(2.5)	(0.9)	(1.4)
CONSTANT	0.45**	0.42**	0.45**	0.43**	0.47**	0.45**	0.51**	0.48**
	(10.9)	(10.3)	(14.9)	(14.2)	(15.3)	(14.8)	(15.9)	(15.3)
Rho	0.94		1.00		0.99		– 0.80*	
	(0.0)		(0.0)		(0.8)		(1.8)	
Pseudo R-square	1.29		1.21		1.24		1.42	
% separation correctly predicted	1.00		0.92		0.95		0.96	
% consolidations correctly predicted	0.29		0.29		0.29		0.46	

[a]Parameter estimates shown with standard error in parentheses below.
* = statistically significant at 0.10;
** = statistically significant at 0.01. Number of observations = 298. Pseudo R-square is that of McIntosh and Dorfman (1992). % separation correctly predicted is the proportion of times the regression correctly predicted that the two cities would remain independent; % consolidations correctly predicted is the proportion of times the regression correctly predicted the two cities would consolidate. Dependent variable is *CONSOLIDATED*.

big and small city encourages the small city to consolidate. Because small cities must agree to any consolidation, it seems that all else constant a consolidation is more likely the more similar in size two cities are. The additional control variable *PROPERTY VALUE* is negative: cities with ample property value tend to not seek consolidation. On the other hand, *DIFFERENCE IN PROPERTY VALUE* is statistically insignificant, a result robust to all specifications.

Demographic differences are expected to discourage consolidation, and the "Add All Demographics" columns of results generally confirm this hypothesis. While differences in educational attainment are insignificant, greater differences in income and racial composition discourage consolidation. The *DIFFERENCE IN PUPILS* result remains robust.

To determine the robustness of the demographic characteristics, and to further confirm the robustness of the size difference results, different combinations of demographic variables are used as explanatory variables. In the next several models in Tables 5 and 6, the parameter estimates of the racial and income difference variables remain nearly the same size, and they are always negative and statistically significant. On the other hand, the parameter estimates of *DIFFERENCE IN EDUCATION* vary from 0.08 to –0.48. They are significant only when *DIFFERENCE IN INCOME* is omitted, suggesting that *DIFFERENCE IN EDUCATION*'s sporadic significance stems from omitted variable bias: it is picking up some of the effect of omitting *DIFFERENCE IN INCOME*. The focus variable, *DIFFERENCE IN PUPILS*, remains robust. A bigger difference in size makes a big city more eager to consolidate with the small city. Economies of scale gains may tempt small cities, as the parameter estimates of the other size variables indicate, but all else constant increasing the size gap discourages small cities from consolidating schooling with larger cities.

The final prediction of Ellingsen's (1998) model that can be investigated is whether excessive consolidation occurs. Recall that Ellingsen predicts that some consolidations that would enhance total welfare will not take place. Only 4% of the pairs of cities that the estimation predicts will remain separate actually consolidated. In contrast, 55% of the cities predicted to consolidate actually retained sole control over schooling provision. This under-consolidation is consistent with Ellingsen's observation that there may be too little consolidation but never excessive consolidation.

VII. CONCLUSION

The theoretical model of Ellingsen (1998) predicts that in a setting in which both cities must agree to consolidate public good provision, differences in size discourage consolidation. Specifically, the model predicts that larger differences in size make big cities more willing to consolidate. But larger differences in size make small cities less willing to consolidate. Because smaller cities have veto power, larger differences in size hinder consolidation.

Ellingsen's model deals with pure and local public goods provided by cities. The current study cannot directly test Ellingsen's model: his model is based on assumptions that do not correspond well to the empirical section. However, Ellingsen's model yields certain predictions, and these predictions are tested using an Ohio data set of public schooling provision. 298 pairs of municipalities are found that could potentially form a consolidated school district, and the determinants of the consolidation decision are examined for the larger and the smaller municipality in each potential consolidation pair.

Bigger differences in size make large cities more likely to prefer consolidation and small cities more likely to prefer separation. This result contrasts with Ferris and Graddy (1988) and Brasington (1999b), who find that bigger and smaller cities are most likely to consolidate, with cities of similar size tending to retain separate service provision. The current study is better designed to investigate the role of size differences because of the way it arranges the data and investigates the decision-making process of the larger and smaller cities separately.

The evidence also supports the conjecture that there will not be excessive consolidation, but that pairs of cities that would seem to benefit by consolidating actually remain separate. The results further suggest that, with the important exception of size differences, big and small cities make their consolidation decisions based on the same criteria. The most influential determinants of the consolidation decision are size and property value factors. Differences in income and racial composition may matter, but the results for educational differences are inconsistent.

Public schooling takes the lion's share of local government expenditures, but there are other important publicly-provided goods. Future work should examine the consolidation of police protection services, fire protection services, and garbage collection services. Such work should investigate the determinants of consolidation from the point of view of the larger and smaller city in each potential consolidation pair. Future work could also examine the determinants of the consolidation of public services from the point of view of the richer and poorer, and the whiter and less white city (Brasington, 2003).

Forming and breaking alliances with each other is not the exclusive domain of local governments. Nations often reunite (Germany, Yemen), separate (Yugoslavia, U.S.S.R.), and jointly provide public goods by joining international organizations (NATO). Their decisions may be influenced by differences in size, demographic characteristics, and preferences over public good provision. Perhaps the lessons and empirical techniques of the current study can prove useful in an international setting.

ACKNOWLEDGMENTS

Thanks to the following people for helpful comments: Chris Wheeler, Manfred Dix, Bill Oakland, Mark Berger, Dek Terrell, Carter Hill, and the referees; all of them are exonerated from remaining errors. Thanks to the University of Kentucky for providing an office during my junior sabbatical, during which much work on the revision of this paper was done. First draft completed August, 2000 at Tulane University.

NOTES

1. Austin (1999) investigates the annexation of suburban land by cities, and Filer and Kenny (1980) investigate the formation of city–county governments.
2. If collective action boosts environmental quality in one school district, property values in neighboring communities may benefit as well (Brasington and Hite, 2002; Leggett and Bockstael, 2000), another example of the spillover benefits of schooling.

3. The investigation of whether large cities provide more aggregate public schooling services than small cities for the sample at hand proceeds as follows. The aggregate amount of public schooling services provided is defined as the product of the number of pupils in the city and the city's taste for schooling services. A city's taste for schooling services is measured by proficiency test passage rates, a measure of schooling outcomes. First, 298 pairs of cities that could consolidate are found. Some of these cities provide their own schooling services, but some have decided to jointly provide schooling with a neighboring city. For cities in consolidated school districts, it is not possible to observe the proficiency test passage rate they would have had if they provided schooling separately, so the proficiency test rate is predicted. The regressions may be found in Brasington (1999b). The aggregate amount of public schooling is calculated using these estimates and the number of pupils. The aggregate amount of public schooling is calculated separately for the larger and the smaller member of each pair of cities that could consolidate. A dummy variable "bigmore" is created, taking the value 1 if the larger member provides a larger aggregate amount of schooling. "Bigmore" has a mean of 0.953, suggesting that over 95% of the time, the bigger city provides a larger amount of aggregate public schooling services.
4. Ellingsen's (1998) model assumes that there is no room for bargaining between jurisdictions. Such an assumption fits the real-world example well because jurisdictions cannot credibly commit to pre-consolidation bargaining. Thanks to a referee for pointing this out.
5. Ohio law is consistent with the theoretical model of Ellingsen (and the empirical technique) in that consolidation only occurs if both communities choose to do so.
6. Excluding municipalities that send their children to more than one school district effectively under-samples consolidated school districts, because the municipality did share schooling services with at least one neighbor. The under-sampling requires the use of choice-based sampling in the empirical estimation.
7. The definitions of central city, suburban, and rural Ohio school districts follow those used in Brasington (1999a).
8. Examples include Callan and Santerre (1990), Andrews et al. (2002), Duncombe and Yinger (2000), Duncombe et al. (1995), Ratcliffe et al. (1990), Lewis and Chakraborty (1996), and Kenny (1982).
9. On the other hand, if a large community is beyond the optimal size diseconomies of scale may dissuade it from consolidating. To investigate this possibility, the optimal size of a school district was estimated following Ratcliffe et al. (1990) and Callan and Santerre (1990). The estimated optimal size exceeds the maximum in the sample, so any consolidation reaps additional scale economies.
10. Brasington (1999b) suggests that racial and income differences may only matter when differences in state consolidation laws are not properly controlled for.
11. Austin (1999) uses a similar estimator based on a bivariate logit. Like Austin, the current study assumes independence between observations. Allowing cross-observational correlation creates formidable convergence problems, even if it would correct the consistency problem discussed later.
12. See note 9.

REFERENCES

ANDREWS, M., DUNCOMBE, W. and YINGER, J. (2002). Revisiting economies of size in American education: Are we any closer to a consensus. *Economics of Education Review*, **21**, 245–62.

AUSTIN, D. A. (1998). A positive model of special district formation. *Regional Science and Urban Economics*, **28**, 103–22.

——— (1999). Politics vs economics: Evidence from municipal annexation. *Journal of Urban Economics*, **45**, 501–32.

Baldwin's Ohio Revised Code, Annotated. (1995). Title 33, Chapter 3311, Sections 3311.53 and 3311.54. New York: Banks-Baldwin Law Publishing Company, division of West Publishing Company, 493–4.

BORCK, R. (1998). Centralization of public good supply with majority voting. *Finanzarchiv*, **55**, 21–40.

BRASINGTON, D. M. (1999a). Central city school administrative policy: Systematically passing undeserving students. *Economics of Education Review*, **18**, 201–12.

BRASINGTON, D. M. (1999b). Joint provision of public goods: The consolidation of school districts. *Journal of Public Economics*, **73**, 373–93.

BRASINGTON, D. M. (2000). Competition in the supply of local public goods: The case of school quality. Tulane University working paper.

BRASINGTON, D. M. (2003). Snobbery, racism, or mutual distaste: What promotes and hinders cooperation in local public good provision? *The Review of Economics and Statistics,* Forthcoming.

BRASINGTON, D. M. and HITE, D. (2002). Demand for environmental quality: A spatial hedonic analysis. Louisiana State University working paper.

CALLAN, S. J. and SANTERRE, R. E. (1990). The production characteristics of local public education: A multiple product and input analysis. *Southern Economic Journal*, **57**, 468–80.

CONLEY, J. and DIX, M. (1999). Optimal and equilibrium membership in clubs in the presence of spillovers. *Journal of Urban Economics*, **46**, 215–29.

DUNCOMBE, W., MINER, J. and RUGGIERO, J. (1995). Potential cost savings from school district consolidation: A case study of New York. *Economics of Education Review*, **14**, 265–84.

DUNCOMBE, W. and YINGER, J. (2000). Does school district consolidation cut costs? Center for Policy Research, Syracuse University working paper.

ELLINGSEN, T. (1998). Externalities vs internalities: A model of political integration. *Journal of Public Economics*, **68**, 251–68.

FERRIS, J. and GRADDY, E. (1988). Production choices for local government services. *Journal of Urban Affairs*, **10**, 273–89.

FILER, J. E. and KENNY, L. W. (1980). Voter reaction to city-county consolidation referenda. *Journal of Law and Economics*, **23**, 179–90.

FISCHEL, W. A. (2002). An economic case against vouchers: Why local public schools are a local public good. Dartmouth College Economics Department Working Paper 01-02, draft of February 2002.

GARASKY, S. and HAURIN, D. R. (1997). Tiebout revisited: Redrawing jurisdictional boundaries. *Journal of Urban Economics*, **42**, 366–76.

GREENE, W. H. (1997). *Econometric Analysis*, 3rd edn. U.S.A: Prentice-Hall, 204.

GUESNERIE, R. and ODDOU, C. (1981). Second best taxation as a game. *Journal of Economic Theory*, **25**, 67–91.

HAVEMAN, R. H. and WOLFE, B. L. (1984). Schooling and economic well-being: The role of nonmarket effects. *Journal of Human Resources*, **19**, 377–407.

HENDERSON, J. V. and THISSE, J. F. (1997). On strategic community development. Centre for Economic Policy Research Discussion Paper.

HOYT, W. H. (1999). Leviathan, local government expenditures, and capitalization. *Regional Science and Urban Economics*, **29**, 155–71.

KENNY, L. W. (1982). Economies of scale in schooling. *Economics of Education Review*, **2**, 1–24.

KENNY, L. W. and SCHMIDT, A. B. (1994). The decline in the number of school districts in the U.S.: 1950-1980. *Public Choice*, **79**, 1–18.

LEGGETT, C. G. and BOCKSTAEL, N. E. (2000). Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management*, **39**, 121–44.

LEWIS, W. C. and CHAKRABORTY, K. (1996). Scale economies in public education. *Journal of Regional Analysis and Policy*, **26**, 23–35.

MCINTOSH, C. S. and DORFMAN, J. H. (1992). Qualitative forecast evaluation: A comparison of two performance measures. *American Journal of Agricultural Economics*, **74**, 209–14.

OHIO DEPARTMENT OF EDUCATION (1985). *Maps of Ohio School Districts: City, Exempted, Local* Columbus, OH: Ohio Department of Education.

OHIO MUNICIPAL ADVISORY COUNCIL (1993). *Compendium of Municipal Debt*. Independence, OH: Ohio Municipal Advisory Council.

PECORINO, P. (1999). The effect of group size on public good provision in a repeated game setting. *Journal of Public Economics*, **72**, 121–34.

POIRIER, D. (1980). Partial observability in bivariate probit models. *Journal of Econometrics*, **12**, 209–17.

RATCLIFFE, K., RIDDLE, B. and YINGER, J. (1990). The fiscal consolidation of school districts in New England: Is small beautiful? *Economics of Education Review*, **9**, 81–99.

STIGLITZ, J. E. (2000). *Economics of the Public Sector*, 3rd edn. New York: W.W. Norton & Company, 426–9.

TIEBOUT, C. (1956). A pure theory of local public expenditure. *Journal of Political Economy*, **64**, 416–24.

WYCKOFF, J. H. (1984). The nonexcludable publicness of primary and secondary public education. *Journal of Public Economics*, **24**, 331–51.