The Political Economy of School District Mergers

Nora Gordon and Brian Knight[*†]

September 28, 2005

Abstract

The number of school districts in the United States has fallen from around 130,000 in 1930 to just under 15,000 at present. Despite this large observed decline, many districts resisted consolidation before ultimately merging and others never merged, choosing to remain at enrollment levels that nearly any education cost function would deem inefficiently small. Why do some districts voluntarily integrate while others remain small, and how do those districts that do merge choose with which of their neighbors to do so? In addresing these questions, we empirically examine the role of potential economies and diseconomies of scale, the loss in autonomy associated with heterogeneity between merger partners, and the role of state governments. We first develop a simulaton-based estimator that is rooted in the economics of matching and thus accounts for three important features of typical merger protocol: two-sided decision making, multiple potential partners, and spatial interdependence. We then apply this methodology to examine the determinants of a wave of school district mergers in the state of Iowa during the 1990s. Preliminary results highlight the importance of economies of scale, diseconomies of scale, and state financial incentives for consolidation.

*Respectively, Department of Economics, University of California, San Diego, email: ngordon@ucsd.edu, and Department of Economics, Brown University, email: Brian_Knight@brown.edu. We thank Julie Cullen, Andrew Foster, Vernon Henderson, Sean Campbell, Kim Sau Chung, Jose Alvaro Rodrigues-Neto and seminar participants at Northwestern, Yale, and Brown for helpful comments. We are grateful to Lynn Carlson for GIS coding, to Xiaoping Wang for assistance with administrative data, and to Guy Ghan for sharing his extensive institutional expertise. Svetla Vitanova provided outstanding research assistance. We are solely responsible for the content of this paper, which does not reflect the views of the Advanced Studies Fellowship program or its funders, the Spencer Foundation and the William and Flora Hewlett Foundation.

† PRELIMINARY AND INCOMPLETE. PLEASE DO NOT CITE OR CIRCULATE WITHOUT PERMISSION OF THE AUTHORS.

1 Introduction

Throughout the twentieth century, bureaucrats, professional educators, and elected officials in the United States encouraged school districts to consolidate. Proponents of consolidation argued that by consolidating, districts would gain from economies of scale: high schools could offer more subjects, elementary schools could separate classes by grade level, and the quality of education could generally be improved at lower costs in larger consolidated schools and districts than in smaller ones. But many school districts resisted: residents consistently voted in favor of retaining their small districts, revealing that they preferred local control over the types of schools their children attended, who their children's classmates would be, and the determination of local tax rates to their own estimation of the potential efficiency gains so touted by consolidation's proponents. Ultimately, many states enacted legislation mandating or providing strong financial incentives for districts to consolidate, prompting sharp drops in the number of school districts (see Hooker and Mueller (1970) for an overview of such legislation), and a vast number of these political battles were resolved in favor of consolidation. As Figure 1 shows, the number of school districts in the United States plummeted from around 130,000 in the early 1920s to just under 15,000 today.

What explains the pattern of consolidations over this period? Why do some districts voluntarily integrate while others choose to remain small? How do districts that do merge choose with which of their neighbors to do so? In attempting to answer these and related questions, a theoretical and empirical literature has investigated the role of several factors.[1] First, regarding the role of population, small districts may benefit from any economies of scale associated with consolidation due to the spreading of fixed costs over more taxpayers. On the other hand, large districts may be discouraged from consolidation due to potential diseconomies of scale. Second, if the potential merger partner has different preferences for publicly-provided goods, the median voter may fear the loss in autonomy associated with consolidation. This heterogeneity in preferences, along with other forms of heterogeneity, may serve as a repelling force. Finally, higher level governments, U.S. states in particular, may either encourage or discourage consolidations through the form of annexation laws or through state aid formulas.

In evaluating the impact of these factors, researchers are confronted with several methodological issues. In particular, standard econometric models of discrete choice fail to account

[1] We survey the relevant empirical literature in the next section. For an overview of the theoretical literature on endogenous borders, see Alesina and Spolaore (1997 and 2003), Bolton and Roland (1997), and Persson and Tabellini (2000).

for three key features of standard merger protocol. First, mergers must typically be approved by voters in both districts, and the decision-making is thus two-sided. Second, in addition to deciding *whether or not* to merge, districts have multiple borders and thus must decide *with whom* to merge. Third, merger decisions are spatially interdependent; if two districts A and B merge, for example, then the choice set is altered for all districts sharing a border with either A or B. While the bivariate Probit model of Poirier (1980) accounts for the first feature and the multinomial logit model accounts for the second feature, we know of no estimators that account for all three of these features of merger protocol.

To overcome these limitations of existing estimators, we first develop an econometric model of discrete choice that accounts for these three key features of the merger protocol. This model is rooted in the economics of one-sided matching and thus allows for two-sided decision making, multiple potential partners, and spatial interdependence. In the context of this model, we show that, under a seemingly reasonable restriction on preferences, which we refer to as *symmetry in match quality*, a unique stable matching exists. Moreover, this stable matching can be calculated via a simple iterative algorithm. Finally, we develop a simulation-based estimator, which uses this iterative algorithm in order to calculate the probability of a merger between any two adjacent districts.

To illustrate its value, we then apply this methodology through an analysis of school district mergers in the state of Iowa, which offered significant financial incentives for mergers during the early 1990s. Due in part to these incentives, over 50 mergers involving more than 100 districts occurred during this period, and, due to these mergers, the number of districts fell from 430 in 1991, the first year included in the analysis, to 371 in 2002, the final year in the analysis. In order to identify all *potential* mergers, which can occur only between adjacent districts, we have obtained a school district map from 1989, just before the start of the sample period. In order to examine the role of district characteristics in these mergers, we have also collected data on pre-merger district characteristics, such as population, demographics, and property values. Finally, in order to examine the role of the state of Iowa, we have calculated the state-level financial incentives specific to each potential merger. Preliminary results demonstrate the importance of economies of scale as well as diseconomies of scale in explaining the patterns of mergers in Iowa during this time period. We also find an important role for the state financial incentives in encouraging these mergers. We find only a minor role for heterogeneity, although these results can not necessarily be generalized to other states.

The paper proceeds as follows. In section 2, we describe the methodology and findings of the existing literature. Section 3 develops the econometric model, which is then applied to school districts mergers in Iowa in Section 4. Finally, section 5 concludes and discusses possible extensions to the methodology.

2 Existing Literature

Several existing empirical studies shed light on the role of factors underlying political integration. Alesina, Baqir, and Hoxby (2004) examine the number of jurisdictions, including school districts, within U.S. counties over the period 1960-1990 and find evidence for a trade-off between economies of scale and heterogeneity in both race and income. That is, counties with high levels of heterogeneity in these dimensions tend to have more school districts, all else equal. On the other hand, they find little effect of heterogeneity in religion or ethnicity. Regarding the role of state governments, the authors find that the strength of annexation laws matter in determining the number of school districts within a state. In a study analyzing the role of state characteristics in determining the number of school districts within a state, Kenny and Schmidt (1994) find that the decline the number of school districts between 1950 and 1980 can be explained by the decline in farming and corresponding increase in population density, the increased importance of state aid, and the increased prominence of teacher unions.

Relative to this literature, which examines changes in the number of school districts within larger geographic units, such as states and counties, we are more focused on individual merger decisions involving adjacent school districts. Our approach thus arguably better accounts for constraints on the availability of potential partners that are imposed by existing boundaries, and variation in these constraints could lead two otherwise identical districts to have different merger patterns. On the other hand, our approach is most appropriate within a single state, while the papers by Kenny and Schmidt (1994) and Alesina, Baqir, and Hoxby (2004) are more naturally suited to an examination of multiple states. Thus, we view our analysis as complementary to this existing line of research.

The only studies of which we are aware that examine the decisions of adjacent school districts to consolidate are a series of papers by Brasington. Brasington (1999) identified 298 pairings of Ohio communities that either do or potentially could jointly provide education services through a single school district. He then estimates a bivariate Probit model developed by Poirier (1980); this model allows for both communities to have veto power over the merger decision and thus a merger is observed only if it is supported by both districts. Using this econometric methodology, he finds that small and large districts tend to jointly provide education services, while medium-sized communities do not enter such arrangements. Neither racial heterogeneity nor income levels explain these patterns. In two follow-up papers, Brasington uses the same dataset from Ohio but allows for the coefficients to vary between the larger and smaller merger partner (Brasington, 2003b), between the richer and poorer district (Brasington, 2003a), and between the more and less white district (Brasington, 2003a).

Relative to these papers by Brasington, our paper provides several contributions. First, while all of Brasington's papers account for the two-sided nature of mergers, they do not account for the two other key features described above: districts must choose from one of several potential partners and merger decisions are spatially interdependent. A failure to account for these features of merger decisions may lead to both specification errors as well as incorrect inference due to the statistical dependence of the observations. Second, while Brasington uses school district characteristics, such as enrollments, test scores, and property values, from the early 1990s to explain consolidation decisions in Ohio, many of which occurred during the 1930s and 1960s, we better model the timing of the merger decisions. The failure to account for these timing considerations could lead to problems in interpretation. For example, if mergers lead to a convergence of school district characteristics, then Brasington's analysis may incorrectly attribute the decision to merge to similarities in district characteristics. Our study, by contrast, measures school district characteristics during the year in which the merger decisions were made, allowing us to separately identify the causes of mergers from their subsequent effects.[2] While we have provided several methodological contributions to this literature, Brasington's specification is somewhat more general in other dimensions. In particular, it allows for a correlation between the unobserved preferences for consolidation between the two merger partners and, in the two follow-up papers, allows the coefficients to vary across the two potential merger partners. Thus, we again view our approach as complementary to this existing line of research.

3 Methodological approach

In analyzing the determinants of mergers between jurisdictions, the analyst is immediately confronted by three methodological challenges. First, in order to take place, mergers must be approved by both districts, and the problem is thus two-sided. Second, in addition to

[2] In separate work (Gordon and Knight, 2005), we are examining the effects of these mergers on subsequent school district fiscal outcomes.

deciding whether or not to merge, districts typically have multiple borders and thus must decide with whom to merge out of this set of potential partners. Finally, merger decisions are spatially interdependent across districts. In order to overcome these challenges, we develop a simulation-based estimator rooted in the economics of matching. We next describe the economic environment and the stability concept before deriving the econometric estimator.

3.1 Matching model

Consider a set of school districts. Districts have an opportunity to merge with other districts sharing a common border. Also, mergers can only occur between two districts, and districts may choose to remain alone. Finally, mergers must be approved by both districts.

This merger environment can be modeled as a one-sided matching game. In particular, a matching is defined as a set of merger assignments; each district is assigned either a single merger partner or is assigned to remain alone. Following the literature on matching, we use stability as the equilibrium concept. A stable matching is a matching in which 1) no district prefers to remain alone over merging with their assigned partner, and 2) no two districts prefer to match with each other over their respective merger assignments.

In one-sided matching situations such as this one, such stable matchings do not exist in general, and when they do exist, are not necessarily unique. Consider, for example, three districts A, B, and C all of which border each other. Suppose that all three districts prefer any merger over remaining alone. Suppose further that A prefers B over C ($B \succ_A C$), B prefers C over A ($C \succ_B A$), and C prefers A over B ($A \succ_C B$). Denote this odd cycle as ABC. In this case, no stable matching exists since any merger can be broken by the unmerged district. On the other hand, with a four-distict case and an even cycle such as $ABCD$, multiple stable matchings may exist.[3]

Such non-existence and multiplicity create severe problems in empirical work. Fortunately, a simple restriction on preferences guarantees both existence and uniqueness. Before introducing such a restriction, define the utility, or gains, to district i from a merger with district j as follows:

$$U_{ji} = A_j + I_i + Q_{ji}$$

[3] That is, if all districts prefer any merger over remaining alone, A merging with B and C with D is a stable matching so long as $A \succ_B D$ or $C \succ_D B$. However, A merging with D and B with C is also a stable matching so long as $D \succ_A C$ or $B \succ_C A$.

where A_j represents the attractiveness of district j as a partner and is valued equally by all of j's potential partners, I_i represents district i's inclination to merge with any of its potential partners, relative to remaining alone, and Q_{ji} represents the quality of the match between districts i and j, as valued by district i.[4] Utility from remaining alone is normalized to zero ($U_{ii} = U_{jj} = 0$). Given our empirical motivation, we assume throughout that districts have strict preferences over their potential merger partners.

While this specific formulation of utility places no restrictions on preferences, we next introduce the restriction of *symmetry in match quality*:

$$Q_{ji} = Q_{ij}$$

That is, conditional on the attractiveness of each district as a partner and the inclination of each district to merge with any of its partners, the quality of the match is equally valued by the two districts. Using this restriction, we have established the following result:

Proposition: Under symmetry in match quality and strict preferences, there exists a unique stable matching.

Proof: See Appendix.

Intuitively, the restriction of symmetry in match quality places enough symmetry on preferences over merger partners in order to rule out the cycles described in the above examples. This in turn guarantees the existence of a unique stable matching.[5]

While Proposition 1 is interesting from a theoretical perspective, its usefulness for empirical work is less obvious. Fortunately, under the assumptions of symmetry in match quality and strict preferences used in the proposition, this stable matching can always be calculated using the following iterative algorithm:

Step A: Match mutual 1st choices (including option to remain alone)

Step B: Remove matched districts from map

Step C: Re-rank from remaining borders and return to Step A

Again, the restriction of symmetry in match quality rules out cycles and thus guarantees that a border consisting of two districts that are mutual first choices can always be found

[4] We do not explicitly model geographic constraints here. However, these constraints can be easily incorporated into preferences by setting $Q_{ji} = -\infty$ for non-adjacent districts.

[5] In independent work, of which we became aware after developing our theoretical results, Rodrigues-Neto (2005) showed that, under symmetric utilities ($U_{ij} = U_{ji}$) and strict preferences, there is always a unique stable matching. While our restriction of symmetric match quality appears to be more general at first glance, these two restrictions turn out to be theoretically equivalent. In particular, if $U_{ji} = A_j + I_i + Q_{ji}$, where $Q_{ji} = Q_{ij}$, then preferences can be represented equivalently by $V_{ji} = U_{ji} - I_i + A_i$ and thus $V_{ji} = V_{ij}$.

in Step A. Our ability to calculate the stable matching via this simple iterative algorithm suggests that a simulation approach may be productive from an econometric perspective. We next turn to the development of such an empirical approach.

3.2 Estimation

Consider an empirical version of the above utility function defined over merger partners:

$$U_{ji} = X_j\theta_x + Z_i\theta_z + f(W_i, W_j)\theta_w + \varepsilon_{ji}$$

where X_j represents observed measures of the attractiveness of district j as a partner, Z_i represents observed measures of district i's inclination to merge with any of its potential partners, relative to remaining alone. The observed quality of the match is given by $f(W_i, W_j)$,while the unobserved quality is given by ε_{ji}; this unobserved match quality is assumed to be distributed type I extreme value and independently across borders. Finally the vector $\theta = (\theta_x, \theta_z, \theta_w)$ represents parameters to be estimated. It is clear that symmetry in match quality is satisfied whenever $f(W_i, W_j) = f(W_j, W_i)$ and $\varepsilon_{ji} = \varepsilon_{ij}$, and we impose these conditions throughout the remainder of the paper.

Given the two-sided nature of the problem, multiple potential partners for each district, and the interdependence of merger decisions, it is clear that no closed form solution exists for the probability of a merger between any two districts. As an alternative to analytically expressing the probability of a merger between any two adjacent districts, one can use the simulation methods for discrete choice models due to Lerman and Manski (1981). In particular, for replication $r = 1, 2, .., R$, an unobserved match quality ($\varepsilon^r_{ji} = \varepsilon^r_{ij}$) can be drawn randomly from the type-I extreme value distribution for each border, and, given a set of parameters (θ), the iterative algorithm described above can be applied in order to calculate the unique stable matching assignments. Unobserved match qualities can then be re-drawn R times, and the proportion of replications in which i and j merge in a stable matching serves as an estimate of the probability of a merger between i and j:

$$\widehat{\Pr(i, j)} = \frac{1}{R}\sum_{r=1}^{R} y^r_{ij}$$

where $y^r_{ij} \in \{0, 1\}$ is a dummy variable indicating a merger between districts i and j in the stable matching associated with simulation r. In practice, however, a smoothed simulator, which calculates the probability of a merger in each replication, is preferred. The average probability across all replications then serves as the estimate of the probability of a merger between i and j. We describe one possible smoothed simulator, which we use in the empirical application, in Appendix 2.

For estimation purposes, we use the method of simulated moments due to McFadden (1989).[6] Under this method, parameters are chosen in order to minimize the distance between the simulated probabilities of merger and the observed merger decisions. Additional details, including the GMM objective function, the optimal weighting matrix for the instruments, and expressions for the variance-covariance matrix, are provided in Appendix 3.

To summarize, estimation via simulation would proceed as follows:

> *Step 0: For each border, independently draw an unobserved match quality* (ε_{ji}) *from the logistic distribution. Do this* R *times and index the replications* $r = 1, 2, ..., R$.
>
> *Step 1: For each of the* R *replications, and given a set of initial parameter values, run the iterative algorithm described above in order to find a stable matching and the associated merger probabilities. The average of this probability across all simulations is the simulated merger probability.*
>
> *Step 2: Choose a new set of parameter values and return to step 1. Repeat until the GMM objective function is minimized.*

Thus, we have developed an econometric model of discrete choice that overcomes the three key limitations of existing econometric models. In particular, by appealing to the economics of matching and the associated stability concept, the approach accounts for the two-sided nature of the merger protocol, allows each district to have an arbitrary number of potential merger partners, and accounts for the interdependence of merger decisions. To illustrate the value of this approach, we next turn to an empirical application of school district mergers in the state of Iowa during the 1990s.

[6] Given the interdependence in merger decisions (if A merges with B, then C cannot merge with A or B) and our reliance on simulation in calculating the probability of mergers, maximum likelihood estimation is problematic. In particular, the likelihood function is defined over all potential *combinations* of merger decisions. Given that, in the empirical application, we have over 1,000 borders, the number of combinations in quite large, and even with a large number of simulation runs, we may not observe every combination of merger decisions. Thus, our simulation procedure would assign probability zero to combinations of mergers not observed in our simulation runs even though every combination of mergers occurs with positive probability in our empirical model (due to the fact that ε_{ji} is unbounded).

4 Empirical Application

We choose to look at the experience of Iowa in the 1990s for several reasons. First, while the state did provide financial incentives for consolidation, the decision to integrate ultimately rested with the school districts themselves. Earlier consolidation waves in other states, by contrast, were often preceded by state or county-level planning of which specific districts were to be targeted for consolidations. Second, concentrating on more recent consolidation activity gives us access to better data on school district finances and the demographics of students and voters. Third, by looking at a period of consolidation beginning just after the 1989 Census was administered, we have access to the initial school district boundaries as geo-coded in the Census TIGER files. We next describe the data before turning to the potential factors influencing merger decisions.

4.1 Data Sources and Variable Definitions

We draw on a number of data sources to compile our district-year level data on Iowa school districts from 1989 to 2001. Our analysis requires data on the timing and composition of school district consolidations, a listing of potential merger partners, and pre-merger characteristics, including demographics, property values, revenues, and expenditures. Demographic data on school districts come from the Census of Population and Housing for 1990 and 2000, and the Common Core of Data. The Census data from 1989 are tabulated at the school district level in the School District Data Book (SDDB), and we use the "Top 100" dataset from the SDDB. For 1999 data, we use the School District Tabulation (STP2) Data, downloaded from the NCES School District Demographic System. These Census data include richer demographic variables than found in the Common Core, including the distribution of adult educational attainment, age, race and ethnicity, and self-reported home values. Because the Census data are available only decennially, we use the Common Core of Data for less refined demographic variables on an annual basis. For the purposes of our analysis, these variables include the number of total students enrolled in public school, enrollment by grade, and enrollment by race and ethnicity. Data on school district revenues and expenditures are from the School District Finance Data (F33) file, available annually in our time period from the fall of 1989 to the fall of 2001. In particular, we use measures of current instructional spending and state and local revenue, and these measures are converted into per-pupil measures using the corresponding enrollment variable. Finally, we have obtained administrative data from Iowa on property value assessments by year and school district; these data are available beginning in 1991.

In order to identify mergers, we have obtained administrative data on school district consolidations from the Iowa Department of Education dating to 1965. These data list the date on which each consolidation goes into effect, the names and Iowa state identification numbers of the districts merging, and the name and Iowa state identification number of the new district formed. In all cases except one, consolidations involved only two districts. One case did involve three districts; given the econometric complications involved with allowing for three-way mergers, we ignore the role of this single three-way merger in the empirical analysis to follow.

In order to identify potential merger partners, we have obtained a map of school districts from 1989 as geo-coded in the Census TIGER files. According to this map, there were 431 districts and 1,211 borders in 1989. Thus, districts had roughly 5 potential merger partners on average. Given the date of the map, our sample is defined over the period 1991 through 2001, the first and last years, respectively, for which we have complete data.

For tractability reasons, our theoretical and econometric framework is purely static in nature. That is, we do not allow districts to consider how a merger today might alter the pool of potential merger partners in the future. Given our use of panel data, however, we must incorporate such changes in potential merger partners in the construction of our dataset. In particular, if two districts A and B merge in year t to form a new district AB, this new district AB now shares borders with all of A's original borders and B's original borders, and we allow for such subsequent mergers between AB and any of these potential merger partners. Empirically, subsequent mergers were rare; there were only two cases in which a school district, as it existed in 1989, went through two consolidations between 1989 and 2001. Given that recently-merged districts may have less desire to merge again, we include in the econometric analysis a dummy variable for whether or not a district has merged in the previous five years.[7]

4.2 Financial Incentives

Financial incentives applied to school districts voting by November 30, 1990 to make their consolidations effective between July 1, 1991 and July 1, 1993. As Figure 2 shows, districts appear to have responded strongly to these time-specific incentives. Beginning in 1966, the start of our administrative data on consolidations, through 1990, there were zero to three

[7] We have also estimated specifications, which yield similar results, with an indicator for any mergers in the past 10 years.

consolidations per year (with 1966 the only year with more than two). In 1991, the first year for which districts received financial bonuses for consolidating, there were four consolidations. This rose to seven consolidations effective in 1992 and twenty in 1993. Interestingly, this was followed by three additional years of higher than average merger activity, even though districts whose consolidations first took effect in these years were not eligible for the financial incentives. We discuss two possible explanations for these post-1993 mergers below.

The financial incentives had two key components, which are summarized in Table 1. The largest incentive for districts to consolidate between 1991 and 1993 was a five-year reduction in their foundation tax rate. During our sample period, the foundation tax rate in Iowa was $5.40 per $1000 of assessed valuation (5.40 mills). By consolidating, districts with enrollments of fewer than 600 students before consolidating could lower their foundation tax rate to 4.40 mills in the first year post-consolidation, increasing by 0.20 mills per year until reaching 5.40 again in the sixth year after consolidation, where it would remain. Throughout this time, the district would receive supplemental state revenue equal to the decrease in local collections, so that the foundation tax reduction essentially transferred funds from state to local taxpayers with no reduction in total revenue available for local education expenditures. To be clear, the enrollment limit is defined separately for each of the two potential merger partners; all property in the post-merger, or unified, district will be eligible for the lower foundation rate if both partners had enrollment below 600 students. For mergers involving one district below 600 students and one district above 600 students, only the property in the district of the smaller partner is eligible for the lower foundation rate. Mergers involving two large districts, those with enrollments in excess of 600 students, were ineligible for these incentives.

To compute the reduction in the foundation tax rate, we use enrollment figures in order to determine whether the district was above or below 600 students as well as administrative data on assessed values. We then compute the present discounted value of the five-year stream of payments using an assumed discount rate of 3 percent, which is roughly the inflation rate during 1991, and, given the stagnant population in the Iowa, an assumed nominal growth rate in housing values of zero. As shown in Figure 3, mergers only occurred during this subsidy period 1991-1993 along borders in which at least one district had enrollments below 600 students, and the vast majority occurred along borders in which both districts had enrollments below 600 students. Taken together with the spike in mergers during this incentive period, as demonstrated in Figure 2, this evidence suggests that districts strongly responded to the financial incentives in place during this period.

The second major incentive is related to the practice of whole grade sharing (WGS). Under WGS, two distinct districts do not merge their finances and thus maintain independent tax bases; instead, two districts divide responsibility over providing education services for particular grades. A common version of WGS involves both districts maintaining their own elementary school, one district having a middle school serving students from both districts, and the other district having a high school serving students from both districts. Iowa had encouraged whole grade sharing by assigning an additional weight to students in whole grade sharing arrangements when making foundation payments to districts. Specifically, students in WGS arrangements counted as 1.1 "regular" students. The Iowa state legislature changed the school finance law to eliminate additional weights for students in WGS arrangements, but allowed school districts consolidating effective 1991-1993 to continue to weight their enrollments according to the proportion of students previously in WGS for five years after merging. This allowed consolidating districts to retain about $200 per pupil per year over a five-year period that they would have lost had they not merged.[8] Many of the districts consolidating had been involved in WGS agreements, suggesting that districts responded to these incentives.[9]

4.3 Heterogeneity Factors

[8] In order to estimate the monetary value of these whole grade sharing incentives, we first estimate the number of students involved in whole-grade sharing by school district. To generate this estimate, we make the simplifying assumption that a district's enrollment, as reported in the district-level files, is equally distributed across all thirteen (including kindergarten) grades. We then multiply this estimated grade-level enrollment by the number of grades in which there is no reported enrollment across all school-level files for the district. This whole-grade sharing enrollment estimate is thus an estimate of the district's gross exported students. We then multiply the number of students involved in whole-grade sharing by $247, which is 10 percent of the foundation payment in 1991, the first year in which the incentives were in place. Finally, we take the present discounted value of the 5-year stream of payments assuming a discount rate of 3 percent and a nominal growth rate in the foundation payment of 4.5 percent, which is roughly the growth rate realized during this period.

[9] Both the foundation tax rate reduction and continued use of supplemental WGS weights gave districts an incentive to consolidate effective 1991-1993. If we view the decision to consolidate as a choice between WGS and consolidation, districts may have chosen WGS over consolidation prior to 1991 because of the supplemental weights. This reason not to consolidate is not valid for mergers effective after 1993 (although they would still receive greater benefits from merging between 1991 and 1993), so may explain why more districts than average consolidated even after the greatest financial incentives were no longer applicable. Another possibility is that the school board had referred the merger to voters by November 30, 1990 but needed more time to build political consensus before voters ultimately approved the merger, albeit without the financial incentives, in subsequent years.

We focus on three measures of heterogeneity: fiscal, demographic, and spatial. These latter two measures are emphasized in the work by Alesina and Spolaore (1997 and 2003) As a baseline measure of fiscal heterogeneity, we use the difference in per-pupil spending on education, adjusted for tax bases, between the two districts. That is, we estimate preferences for education by dividing per-pupil expenditures, using instructional spending and enrollments in the Census data, by housing values in the district, as self-reported by residents in Census data. To create a measure of heterogeneity, we then take the absolute difference in the measures between the two adjacent districts. To capture potential heterogeneity in tax bases, we also include a measure of the absolute difference in housing prices. For our demographic heterogeneity measures, we examine the absolute difference in percent white among students in the two districts and the absolute difference in percent of adults who have completed high school in the two districts. Finally, regarding spatial heterogeneity, we control for the size of the district, as measured in square miles, as well as the interaction between the sizes of the two districts. If transportation costs are convex in distance, two geographically large districts may have a disincentive to merge with each other.

4.4 Scale Factors

We are also interested in examining the role of economies and diseconomies of scale in these merger decisions. Let $c(N)$ denote the average cost of providing education services to N students. From the perspective of district i, the efficiency gains, or potentially losses, from a merger with district j can be expressed as:

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

For efficiency enhancing mergers $[c(N_i+N_j)<c(N_i)]$, our measure of efficiency gains will be positive. On the other hand, if $c(N_i+N_j)>c(N_i)$, our measure will be negative, suggesting efficiency losses. In terms of an empirical specification, we use the following average cost specification:

$$c(N)=N^{\beta+\gamma N}$$

This specification allows for a wide range of shapes for the cost curve, and, if $\beta>0$ but $\gamma<0$, the cost-curve will be U-shaped. Inserting this specification into our measure of efficiency gains, we have that:

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

Thus, our estimate of β can be considered an estimate of the role of economies of scale in merger decisions, while our estimate of γ can be considered a corresponding estimate of the role of diseconomies of scale.

4.5 Results

Table 2 provides summary statistics at the level of a school district border for our key variables in the econometric analysis. As shown, mergers were more likely to occur along borders that were eligible for the merger incentives. Given the complexity of interpreting the economies of scale measures, we defer their discussion until the econometric analysis. Regarding heterogeneity measures, mergers were more likely to occur along borders with less heterogeneity in housing values and in smaller districts, as measured by square miles and the interaction between the square miles in the two districts. We next turn to a more formal econometric test of our hypotheses.

Table 3 provides the results from our simulation estimator. The three columns provide results incorporating several alternative measures of heterogeneity. As shown, all three of these specifications report a positive effect of the merger incentives on the propensity to merge, providing evidence that is consistent with the suggestive evidence in figures 2 and 3. As expected, the economies of scale coefficient is negative in two out of three specifications while the diseconomies of scale estimate is positive in all specifications. In order to aid in the interpretation of these results, Figure 4 plots the log cost curve implied by the coefficients in column 1 against district enrollments. Recall that our assumed cost curve is given by $c(N) = N^{\beta+\gamma N}$, and thus we can write the log cost curve as follows:

$$\ln c(N) = (\beta + \gamma N)\ln(N) \tag{1}$$

As shown in Figure 4, these coefficients imply that average costs are minimized at enrollments of about 500 students. Thus, among equally sized districts, the most efficient mergers involve those with enrollments of about 250 each, and mergers involving larger districts would entail diseconomies of scale. It is important to note that these estimates of economies of scale and diseconomies of scale should be interpreted as those perceived by the voters when deciding whether or not to integrate. These revealed preference estimates may differ substantially from the economies of scale actually realized by districts through consolidation. Indeed, estimates of education cost functions, as summarized by Andrews, Duncombe, and Yinger (2002), imply that diseconomies of scale may not set in until enrollments reach 6,000 students, although, as the authors point out, this optimal size may be significantly lower in sparsely populated states, such as Iowa, due to transportation costs.

While the results report a consistent story regarding the role of state incentives and district size, the results regarding the heterogeneity measures are more mixed. The results in the first column suggest that fiscal heterogeneity reduces the propensity to merge, while, as shown in column 2, demographic heterogeneity actually seems to increase the propensity to merge. Regarding the spatial heterogeneity measures, the estimates suggest that larger districts prefer to merge. The interaction coefficient, on the other hand, is negative, suggesting that two large districts have a disincentive to merge, presumably due to the significant increases in transportation times associated with such mergers.

5 Conclusion

In this paper, we develop an empirical approach to the study of school district mergers. This method is rooted in the economics of matching and thus overcomes several methodological problems with existing estimators. In particular, our approach allows for two-sided decision making, multiple potential merger partners for each district, and spatial interdependence in merger decisions. While the model does not generate an analytic expression for the probability of a merger, we show that the model can be estimated via simulation techniques. Applying this method to a spate of school district mergers in Iowa during the 1990s, preliminary results demonstrate the importance of state subsidies, economies of scale as well as diseconomies of scale in explaining the patterns of mergers in Iowa during this time period; the results regarding heterogeneity, by contrast, are more mixed. One caveat is that these results, such as our finding that racial heterogeneity played only a minor role, may not generalize to other states and time periods. Iowa has very little racial heterogeneity, and, as noted above, other studies, such as Alesina, Baqir, and Hoxby (2004), have found a strong role for such heterogeneity in terms of predicting the number of school districts within U.S. counties.

Appendix 1: Proof of Proposition

Chung (2000) has shown that no odd cycles implies existence of a stable matching. The first part of our proof shows that, if there are two distinct stable matchings and strict preferences, then a cycle can be created. The second part of the proofs show that under the restriction of *symmetry in match quality* and strict preferences, there are no cycles. Thus, under the *symmetry in match quality* and strict preferences, a unique stable matching exists.

Claim: If there are two distinct stable matchings and strict preferences, then a cycle can be created.

Suppose there are $N > 2$ districts and two distinct stable matchings (A and B). In order for A and B to be distinct matchings, at least one district must be paired with different partners in A and B. Without loss of generality, denote this district as 1 and the partner in A as 2 and the partner in B as 4. Again, without loss of generality, assume that 1 prefers 2 over 4 ($2 \succ_1 4$). Now, denote district 3 as district 2's partner in matching B. In order for matching B to be stable it must the be case that 2 prefers 3 over 1 ($3 \succ_2 1$). In matching A, district 3 must either merge with 4 or a new district, say district 5. If 3 merges with 4, it must be that 3 prefers 4 over 2 in order for A to be stable ($4 \succ_3 2$). But, in order for matching B to be stable, it must be that 4 prefers 1 over 3 ($1 \succ_4 3$) and we thus have that ($2 \succ_1 4$), ($3 \succ_2 1$), ($4 \succ_3 2$), ($1 \succ_4 3$), which we refer to as the cycle 1234. On the other hand, if district 3 mergers with district 5 in matching A, it must be the case that 3 prefers 5 over 2 ($5 \succ_3 2$) in order for A to be stable. Denote 6 as 5's partner in matching B. We thus know that 5 prefers 6 over 3 $6 \succ_5 3$). Now, in matching A, 6 must merge with district 4 or a new district 7. If 6 merges with 4, it must be that 6 prefers 4 over 5 ($4 \succ_6 5$) in order for A to be stable. But, in order for B to be stable, 4 must prefer 1 over 6 ($1 \succ_4 6$) and we have the cycle 123564. On the other hand, if 6 merges with 7, etc. It is thus clear that, given a finite number of districts, this process will eventually lead to a cycle. Thus, if there are two distinct stable matchings and strict preferences, then a cycle can be created.

Claim: Under the restriction of *symmetry in match quality* and strict preferences, there are no cycles.

Suppose not and let the cycle of size C be given by $123...C$. Then, we know that the following preferences hold:

$$\begin{aligned} U_{2,1} &> U_{C,1} \\ U_{3,2} &> U_{1,2} \\ U_{4,3} &> U_{2,3} \\ &... \\ U_{C,C-1} &> U_{C-2,C-1} \\ U_{1,C} &> U_{C-1,C} \end{aligned}$$

Inserting our specification and using the assumption that $Q_{i,j} = Q_{j,i}$, we have that:

$$\begin{aligned} A_2 + Q_{1,2} &> A_C + Q_{1,C} \\ A_3 + Q_{23} &> A_1 + Q_{12} \\ A_4 + Q_{34} &> A_2 + Q_{23} \\ &... \\ A_C + Q_{C,C-1} &> A_{C-2} + Q_{C-1,C-2} \\ A_1 + Q_{1,C} &> A_{C-1} + Q_{C,C-1} \end{aligned}$$

Summing across these conditions, it is clear that the left hand side and right hand side are identical. Hence, a contradiction and no cycle.

Appendix 2: Smooth simulator

For each simulation r, the probability of a merger between two districts i and j can be expressed as the probability of deviations from the stable matching. In particular, denote U_i^* and U_j^* as the utility for districts i and j under the stable matching. After calculating these utilities, we provide each district a small amount of additional information ($\tau\epsilon_{ij}$),which is also distributed type-I extreme value, regarding each of their options. The parameter τ is referred to as the smoothing parameter. For two bordering districts i and j that are not merged together under the stable matching, we can then calculate the probability of a deviation as follows:

$$\begin{aligned}\Pr(\text{deviation}_{ij}) &= \Pr(U_{ij} + \tau\epsilon_{ij} > U_j^* + \tau\epsilon_j, U_{ji} + \tau\epsilon_{ij} > U_i^* + \tau\epsilon_i \\ &= \frac{1}{1 + \exp[(U_j^* - U_{ij})/\tau] + \exp[(U_i^* - U_{ji})/\tau]}\end{aligned}$$

where τ is the smoothing parameter and is chosen to be small; as the smoothing parameter converges to zero, the probability of a merger approaches 0. Thus, in the limit, the smooth simulator approaches the frequency simulator. But, for any positive τ, the probabilities are bounded between zero and one. This simulated probability can thus be interpreted as the probability of a deviation, allowing districts to make mistakes, where the magnitude of the mistakes depends upon the smoothing parameter τ.

For two bordering districts i and j that are merged together under the stable matching, the probability of no deviation is given as follows:

$$\Pr(\text{no deviation}) = 1 - \sum_{k \in B_i} \Pr(\text{deviation}_{ik}) - \sum_{l \in B_j} \Pr(\text{deviation}_{jl}) - \frac{1}{1 + \exp[U_i^*/\tau]} - \frac{1}{1 + \exp[U_j^*/\tau]}$$

where B_i is the set of districts that border district i, other than j, and B_j is a similar set for district j, $\frac{1}{1+\exp[U_i^*/\tau]}$ is the probability that district i would prefer to remain alone over merging with district j and similarly for $\frac{1}{1+\exp[U_j^*/\tau]}$.

Appendix 3: GMM Estimator and Inference

For estimation purposes, we use a simulated method of moments approach, where the objective function is defined below:

$$[I'(y-p)]'W[I'(y-p)]$$

where y is an $N \times 1$ vector of observed merger indicator variables, p is an $N \times 1$ vector of simulated merger probabilities, and I is a $N \times k$ matrix of instruments, or exogenous variables. Finally, W is a $k \times k$ weighting matrix. The optimal weighting matrix is given by the inverse of the variance-covariance matrix:

$$\begin{aligned} W &= var[I'(y-p)]^{-1} \\ &= [I'var(y-p)I]^{-1} \\ &= [I'(E(yy') - pp')I]^{-1} \end{aligned}$$

Note that $E(yy')$ is not necessarily diagonal due to the interdependence in merger decisions. However, we can estimate this matrix via our simulation approach as follows:

$$E(yy') = (1/R)\sum_{r=1}^{R} y^r y^{r\prime}$$

Let $m \geq k$ denote the number of parameters in the vector θ. Then, we calculate the standard errors according to the following variance-covariance matrix:

$$Var(\theta) = (1/N)I'dpWdp'I$$

where $dp = dp/d\theta$.

References

[1] Alesina, Alberto and Enrico Spolaore, 2003, The Size of Nations, MIT Press: Cambridge.

[2] Alesina, Alberto and Enrico Spolaore, 1997, "On the Size and Number of Nations," *Quarterly Journal of Economics*, 112: 1027-1056.

[3] Alesina, Alberto and Reza Baqir and Caroline Hoxby, 2004, "Political Jurisdictions in Heterogenous Communities," *Journal of Political Economy*, 112: 348-396.

[4] Andrews, Matthew and William Duncombe, and John Yinger, 2002, "Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus?," *Economics of Education Review*, 21: 245-262.

[5] Bolton, Patrick, and Gerald Roland, 1997, "The Breakup of Nations: A Political Economy Analysis." *Quarterly Journal of Economics*, 112: 1057-1090.

[6] Brasington, David. 1999. "Joint Provision of Public Goods: the Consolidation of School Districts." *Journal of Public Economics*, 73: 373-393.

[7] Brasington, David. 2003a. "Snobbery, Racism, or Mutual Distaste: What Promotes and Hinders Cooperation in Local Public-Good Provision?" *Review of Economics and Statistics*, 85: 874-883.

[8] Brasington, David. 2003b. "Size and School District Consolidation: Do Opposites Attract?." *Economica*, 70: 673-690.

[9] Chung, Kim Sau. 2000. "On the Existence of Stable Roomate Matchings," *Games and Economic Behavior*, 33(2): 206-230.

[10] Gordon, Nora and Brian Knight, 2005. "The Fiscal Impact of Political Integration: Evidence from School District Consolidations" *mimeo.*

[11] Hooker, Clifford and Van Mueller, 1970, The Relationship of School District Reorganization to State Aid Distribution Systems, National Education Finance Project Special Study No. 11.

[12] Kenny, Larry and Amy Schmidt, 1994. "The Decline in the Number of School Districts in the U.S.: 1950-1980" *Public Choice*, 79: 1-18.

[13] Lerman, R. and C. Manski, 1981. "On the Use of Simulated Frequencies to Approximate Choice Probabilities" in C. Manski and D. McFadden, eds., Structural Analysis of Discrete Data with Econometric Applications, Cambridge: MIT Press, 1981.

[14] McFadden, Daniel, 1989, "A Method of Simulated Moments for Estimation of of Discrete Response Models Without Numerical Integration" *Econometrica*, 57: 995-1026.

[15] Persson, Torsten and Guido Tabellini, 2000, Political Economics: Explaining Economic Policy, MIT Press: Cambridge.

[16] Poirer, Dale. 1980. "Partial Observability in Bivariate Probit Models." *Journal of Econometrics*, 12: 209-217.

[17] Rodrigues-Neto, Jose Alvaro. 2005. "Unique Stable Roomates and Solitaire." forthcoming, *Journal of Economic Theory*.

Table 1: Summary of Merger Incentives

	WGS pre-91	WGS post-91	Reorganize pre-91 or post-93	Reorganize 91-93 (for 5 yrs after reorganization)
Foundation tax rate (pay this to state on district assessed prop valuation)	=$5.40 per $1000	=$5.40 per $1000	=$5.40 per $1000	1st yr = 4.4 2nd yr = 4.6 3rd yr = 4.8 4th yr = 5.0 5th yr = 5.2
Foundation payments to district from state per pupil = F	=F for student in schl in district =1.1F for student in WGS	=F for student in schl in district =F for student in WGS	=F for student in schl in district =F for student previously in WGS	=F for student in schl in district =1.1F for student previously in WGS

Table 2: Summary Statistics for Key Variables

Observation = border / year

mean (standard deviation)

	merger n=51	no merger n=11,620
merger incentive	2.3222	0.4114
(mills, pdv)	(2.3783)	(1.0055)
economies of scale	-0.7892	-0.8468
ln(n1)-ln(n1+n2)	(0.4833)	(0.6307)
diseconomies of scale	-3195.82	-12243.46
n1*ln(n1)-(n1+n2)*ln(n1+n2)	(2067.28)	(29077.94)
spending heterogeneity	0.0400	0.0469
	(0.0473)	(0.0520)
house price heterogeneity	6.8451	10.8776
	(5.2644)	(9.3013)
racial heterogeneity	0.0193	0.0232
	(0.0194)	(0.0334)
heterogeneity in percent	0.0433	0.0454
college graduates	(0.0341)	(0.0350)
square miles	96.6290	133.6089
	38.8276	(68.1594)
square miles *	8694.33	18587.67
neighbor square miles	(3470.56)	(14904.79)
merger in last 5 years	0.0098	0.0357
	(0.0990)	(0.1856)

Table 3: Determinants of School District Consolidations

variable	column 1	column 2	column 3
merger incentive	0.4306	0.4404	0.4418
	0.0008	0.0007	0.0008
economies of scale	-1.3556	0.0271	-1.1572
ln(Ni)-ln(Ni+Nj)	0.0139	0.0046	0.0073
diseconomies of scale	0.0004	0.0003	0.0001
Ni*ln(Ni)-(Ni+Nj)*ln(Ni+Nj)	0.0000	0.0000	0.0000
spending heterogeneity	-7.7809		
	0.0372		
house price heterogeneity	-0.0011		
	0.0002		
racial heterogeneity		7.8851	
		0.0473	
heterogeneity in percent		0.7619	
college graduates		0.0359	
square miles			0.0197
			0.0003
square miles *			-0.0002
neighbor square miles			0.0000
merger in last 5 years	-0.5325	0.9025	-0.8342
	0.0107	0.0419	0.0103
constant	-8.5293	-7.5130	-6.8255
	0.0031	0.0053	0.0069
Observations	11515	11515	11515

Standard errors below coefficients

Figure 1: School districts in the US over time

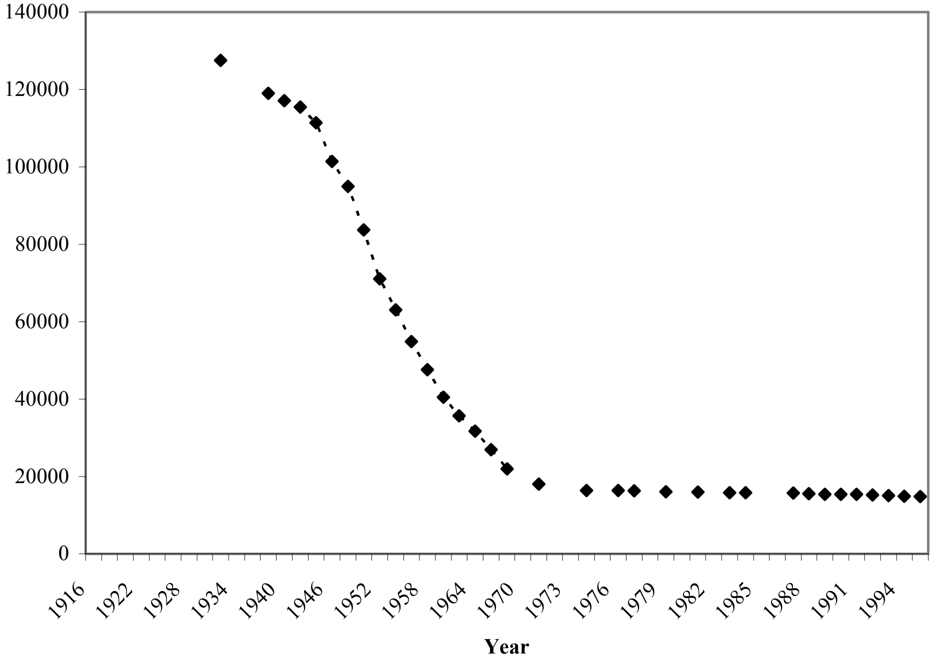
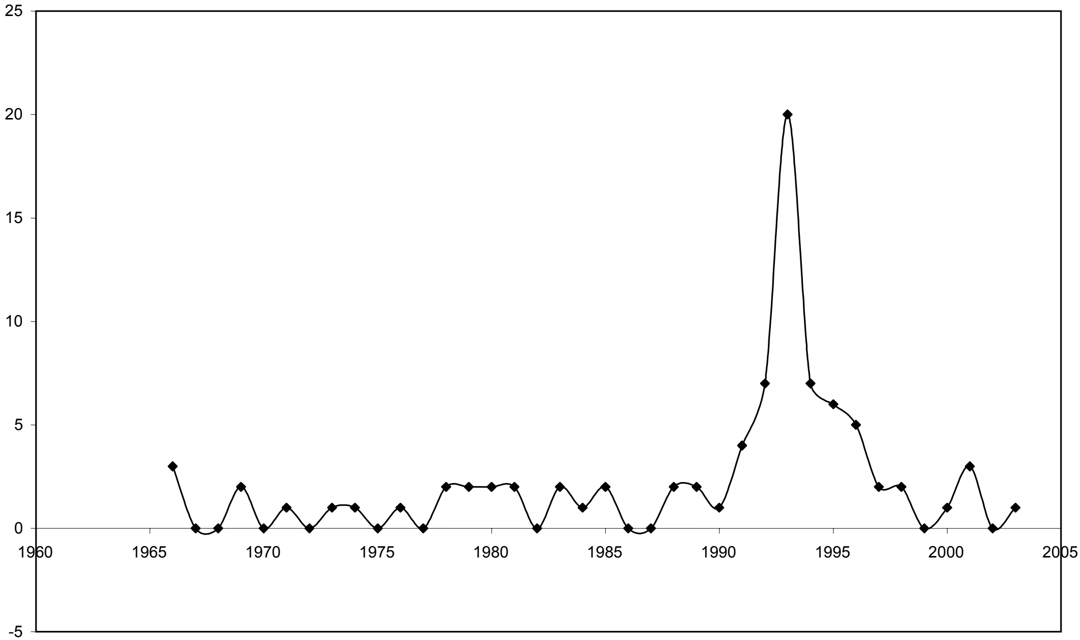


Figure 2: School district consolidations in Iowa, 1966-2003



Year

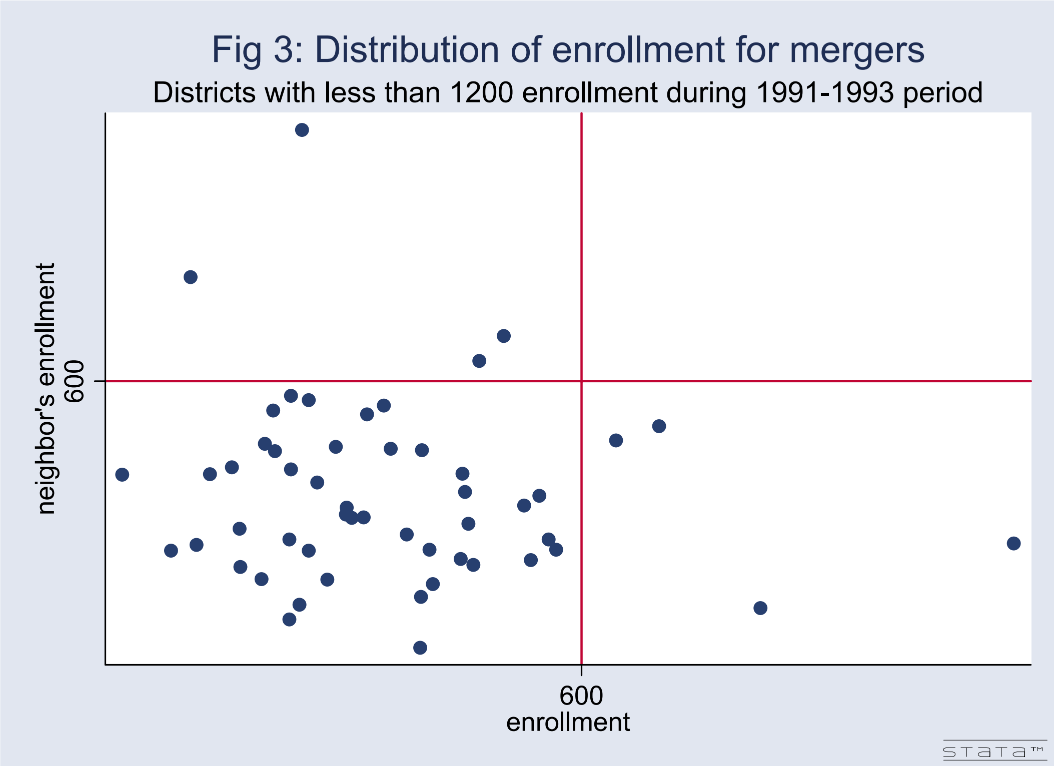


Fig 3: Distribution of enrollment for mergers
Districts with less than 1200 enrollment during 1991-1993 period
neighbor's enrollment
600
600
enrollment
stata™

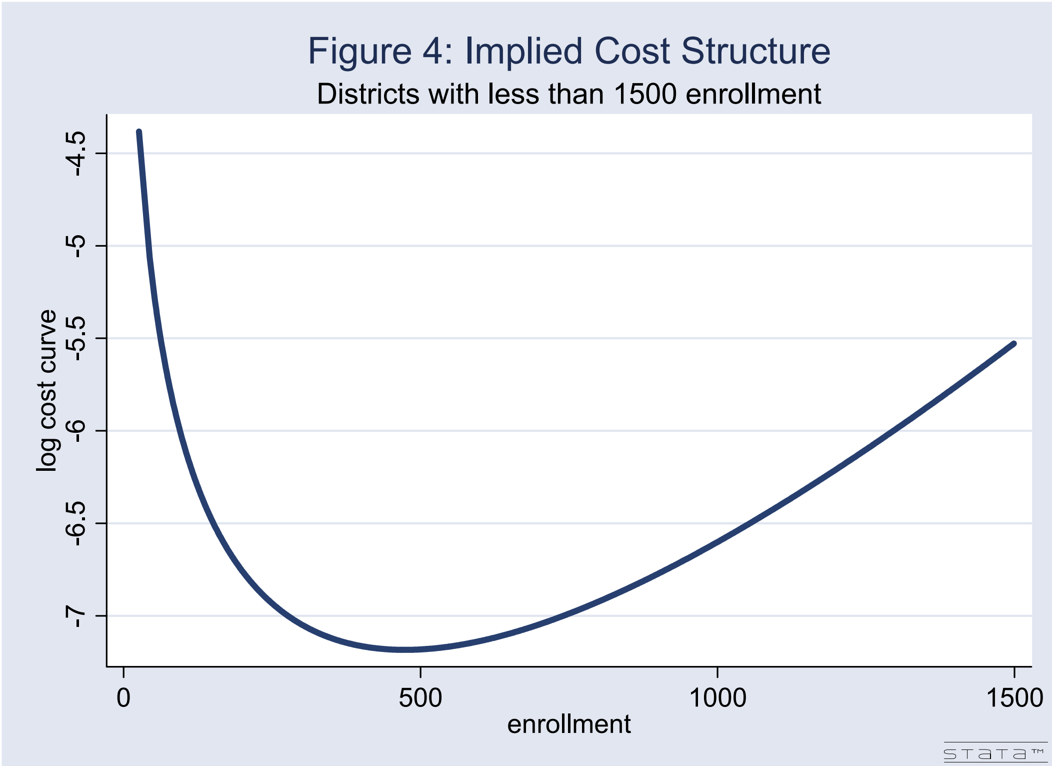


Figure 4: Implied Cost Structure
Districts with less than 1500 enrollment
log cost curve
-4.5
-5
-5.5
-6
-6.5
-7
0
500
1000
1500
enrollment