



A spatial merger estimator with an application to school district consolidation [☆]

Nora Gordon ^a, Brian Knight ^{b,*}

^a Department of Economics, University of California, San Diego, United States

^b Department of Economics, Brown University, United States

ARTICLE INFO

Article history:

Received 27 June 2008

Received in revised form 23 December 2008

Accepted 13 February 2009

Available online 1 March 2009

Keywords:

Local public finance

Fiscal federalism

School district consolidation

Matching games

ABSTRACT

This paper develops a spatial merger estimator to explain political integration generally and then applies this method to a wave of school district mergers in the state of Iowa during the 1990s. Our estimator 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. Rather than simply explaining when a particular region is likely to experience a wave of political integration, our method allows us to explore the factors driving which specific subregional mergers take place. This allows us to explore how those districts that merge choose with which of their neighbors to do so. Our results highlight the importance of state financial incentives for consolidation, economies of scale, diseconomies of scale, and a variety of heterogeneity measures in this particular application. We also demonstrate the power of our estimator, relative to existing estimators, to detect a statistically significant role for heterogeneity factors. While our application is limited to school district consolidation, our method can be adapted to include the salient features of many spatial integration problems.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Economists have made considerable progress in the analysis of matching problems. In some applications, such as in the choice of roommates, there are few restrictions on who can match with whom. In many applications, however, spatial considerations place important restrictions on the set of potential partners. The political integration of contiguous jurisdictions, such as countries or school districts, are examples of such spatial matching problems. The annexation of suburbs by neighboring cities is another prominent example. Similar issues arise when considering mergers of firms for whom geographic location is an important characteristic. Hospitals (see [Dranove and White, 1994](#)) and real estate multiple listing services are two examples of industries with recent merger activity and in which the mergers are strongly influenced by firm location. In this paper, we develop a method for analyzing patterns of spatial mergers. To provide an example of the methodology,

we examine school district consolidation activity in Iowa in the 1990s. As [Fig. 1](#) shows, the number of school districts in the United States has declined precipitously over the twentieth century, so the application to school district mergers is a historically salient one. The method could, however, be applied in any spatial merger context for which the researcher has access to the complete map of jurisdictional borders or the boundaries of firm-level territories.

The theoretical literature on endogenous borders has flourished: [Alesina and Spolaore \(1997, 2003\)](#), [Bolton and Roland \(1997\)](#), and [Persson and Tabellini \(2000\)](#) have focused on the role of potential cost-savings associated with integration as well as the role of heterogeneity in discouraging integration. There is less empirical work, however, examining why political integration occurs in some cases but not others. One reason for the slow progress in this area is the lack of econometric models of jurisdictional merger decisions, reflecting methodological challenges associated with three standard features of merger protocol. First, mergers must typically be approved by voters in both districts (that is, there is two-sided decision making); standard discrete choice models, such as the logit, are designed for single agent decision making. Second, in addition to deciding *whether or not* to merge, districts typically have multiple borders and thus must decide *with whom* to merge. Third, merger decisions are spatially interdependent. That is, if two districts 1 and 2 merge, then the choice set is altered for all districts sharing a border with either 1 or 2. 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 simultaneously account for all three of these features of merger protocol. These three features are all relevant in the institutional

[☆] We thank two anonymous referees, Andrew Foster, Vernon Henderson, Sean Campbell, Kim Sau Chung, Jose Alvaro Rodrigues-Neto, Hal White, and seminar participants at several universities for helpful comments. We are grateful to Svetla Vitanova for research assistance, Lynn Carlson for GIS coding, Xiaoping Wang for assistance with administrative data, and Guy Ghan for sharing his extensive institutional expertise. Gordon thanks the Advanced Studies Fellowship program, and Knight's research was supported by a Salomon Award. 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.

* Corresponding author.

E-mail addresses: negordon@ucsd.edu (N. Gordon), Brian_Knight@brown.edu (B. Knight).

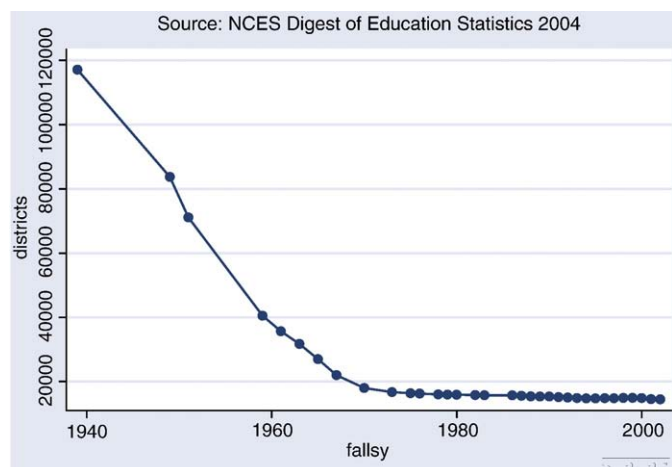


Fig. 1. Number of school districts over time.

context of school district consolidation; they are likely to be relevant, if in weaker forms, for nearly all spatial merger applications.

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. We model this merger environment as a matching game in which jurisdictions choose a partner from the set of adjacent districts, and our approach thus allows for two-sided decision making, multiple potential partners, and spatial interdependence. While existence and uniqueness of equilibrium are not generally guaranteed in such models, 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 we refer to as a spatial merger estimator, which uses this iterative algorithm in order to calculate the probability of a merger between any two adjacent districts in stable matchings.

To illustrate the value of the spatial merger estimator, we then apply this methodology through an analysis of school district mergers in Iowa during the 1990s. Over 50 mergers involving more than 100 districts occurred during this period (see Fig. 2), 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. Our findings highlight the role of potential factors in these merger decisions. First, regarding the role of size, small districts are much more likely to merge, suggesting that they 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 experience diseconomies of scale. Second, we find that like districts are more likely to merge, suggesting an important role for heterogeneity. We also find an important role for state financial incentives in encouraging these mergers. Finally, we show that our estimator is significantly better than existing methods in terms of detecting the role of heterogeneity in driving merger decisions.

The paper proceeds as follows. In Section 2, we describe the methodology and findings of the existing literature. Sections 3 and 4 develop the theoretical and econometric framework, which is then applied to school districts mergers in Iowa in Section 5. Finally, Section 6 concludes.

2. Existing literature

Several existing empirical studies shed light on the role of factors underlying political integration. A first literature examines the incidence of mergers without focusing on the identity of merger partners. Nelson (1990) shows that both income heterogeneity of the

local population and permissiveness of state-level regulations towards local government are correlated with greater numbers of local jurisdictions within a metropolitan area. Alesina et al. (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 in 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 the number of school districts within larger geographic units, such as states and counties, we are focused on specific 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; variation in these constraints could lead two otherwise identical districts to have different merger patterns. Given that we are able to account for how state-specific school finance regimes would affect potential mergers and given the computational difficulty of predicting interdependent mergers for a large number of jurisdictions, our approach is most appropriate within a single state. The reduced-form methodologies of Kenny and Schmidt (1994) and Alesina et al. (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 identities of merger partners, as captured by the decisions of adjacent school districts to consolidate, are a series of papers by Brasington. Brasington (1999) identifies 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 fits a specification that is quadratic in students and finds that small and large districts were most likely to merge, while medium-sized communities do not enter such arrangements.¹ 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 potential merger partner (Brasington, 2003b), between the richer and poorer community (Brasington, 2003a), and between the more and less white community (Brasington, 2003a).

Relative to these papers by Brasington, our method provides several contributions. First and most importantly, 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 only one merger partner from several potential partners, and merger decisions are spatially interdependent.² A failure to account for these features of merger decisions may lead to specification errors. For example, with only two-way mergers and three districts 1, 2, and 3, there are three possible mergers: (1,2), (1,3), and (2,3). Nothing restricts Brasington's approach from internal inconsistencies such as Pr

¹ Note that this quadratic specification includes only own-district students and thus, as will be seen below, does not violate our symmetry assumption.

² To be clear, Brasington's analysis does account for each district having multiple borders, and he does correctly account for all of the possible pairwise combinations.

$(1,2) + \text{Pr}(1,3) + \text{Pr}(2,3) > 1$. In addition, if the protocol allows for a three-way merger, nothing restricts his approach from setting $\text{Pr}(1,2) = \text{Pr}(1,3) = 1$ but $\text{Pr}(2,3) < 1$. Similarly, this failure to account for multiple potential partners and spatial interdependence may lead to incorrect inference due to statistical dependence across borders. By contrast, our approach, which focuses on two-way mergers, accounts for these internal consistency requirements (e.g. $\text{Pr}(1,2) + \text{Pr}(1,3) + \text{Pr}(2,3) \leq 1$) and also accounts for statistical dependence.

While an empirical researcher may be willing to abstract from these specification and statistical issues, a second and related advantage of our approach is the ability to predict a comprehensive map following a change in the merger environment. This feature of our approach may be useful, for example, to a state official who wants to predict how the number and composition of school districts may change following the introduction of a set of merger incentives. Due to the problems of internal inconsistency described above, Brasington's approach cannot predict such a comprehensive map.

Finally, while Brasington uses school district characteristics, such as enrollment, 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 model the timing of the merger decisions as dependent on pre-existing rather than future district characteristics. The failure to account for these timing considerations could lead to problems in interpretation. For example, if district characteristics tend to converge post-merger, then Brasington's analysis may incorrectly interpret similarities in district characteristics between merger partners to a preference for homogeneity, rather than the true source of these similarities: the ex-post convergence in district characteristics. In our empirical application, by contrast, we measure 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.³ 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 an imperfect 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.

3. Theoretical framework

In empirically analyzing the determinants of mergers between jurisdictions, the analyst is immediately confronted by three methodological challenges. Mergers must be approved by both districts, districts typically have more than one potential partner from which to choose, and merger decisions are spatially interdependent across districts. In this section, we develop a simulation-based estimator that is rooted in the economics of matching and thus overcomes these three methodological challenges. We first describe the matching environment and the associated equilibrium concept of stability before deriving the econometric estimator in the next section.

3.1. Matching model

Consider a set of school districts and the following merger protocol. First, mergers can occur only between two adjacent districts, and, for reasons of tractability, we rule out mergers involving three or more districts.⁴ Second, we do not allow for one district to dissolve into multiple districts, again for reasons of tractability.⁵ Third, mergers must be approved by voters in both districts, and the decision-making

is thus two-sided. Fourth, districts may choose not to merge with any adjacent districts; that is, districts may remain unmerged. Finally, given our empirical motivation, we assume that districts have strict preferences and are thus not indifferent between their potential merger partners. The role of these and other key assumptions will be described more completely at the end of this section.

This merger environment can be modeled as a one-sided matching game.⁶ In particular, a matching is defined as a set of merger assignments, in which 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 set of merger assignments in which 1) no district prefers to remain unmerged over merging with their assigned partner, and 2) no two districts prefer to merge with each other over their respective merger assignments, which may include remaining unmerged.

Unfortunately, in one-sided matching situations such as this one, stable matchings do not generally exist, and when they do exist, are not necessarily unique (Roth and Sotomayor, 1990). Consider, for example, three districts 1, 2 and 3 all of which border each other. Suppose that all three districts prefer any merger over remaining unmerged and further that district 1 prefers 2 over 3 ($2 >_1 3$), 2 prefers 3 over 1 ($3 >_2 1$), and 3 prefers 1 over 2 ($1 >_3 2$), and denote this cycle as 123; we subsequently refer to such cycles involving an odd number of districts as an odd cycle. In this case, no stable matching exists since any merger between two districts can be broken by the unmerged district. On the other hand, with a four-district case and an even cycle such as 1234, multiple stable matchings may exist.⁷

Given our objective of developing an empirical methodology, which requires a comparison of mergers observed in the data to those predicted by the econometric model, the problems of non-existence and multiplicity clearly create significant hurdles. Fortunately, a simple restriction on preferences guarantees both existence and uniqueness. Before introducing such a restriction, we define the utility gains to district i from a merger with district j as follows:

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

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 unmerged, and Q_{ji} represents the quality of the match between districts i and j , as valued by district i .⁸ The overall incidence of mergers will be driven by attractiveness and inclination whereas the identity of merger partners will be driven by the quality of the match. Utility from remaining unmerged is normalized to zero ($U_{ii} = U_{jj} = 0$). As noted above, we assume throughout that districts have strict preferences over their potential merger partners.

It should be clear that this specific formulation of utility places no restrictions on preferences, as we can always manipulate the qualities of the match to generate the cycles described above. To eliminate this

³ In separate work (Gordon and Knight, 2007), we examine the effects of these mergers on subsequent school district fiscal outcomes.

⁴ In our data, all mergers except one involve only two districts in any given year.

⁵ In our data, two districts are dissolved over the full time period. These districts are small and their enrollment is dispersed among pre-existing districts, so no new jurisdictions are formed.

⁶ In a one-sided matching game, also known as the roommates problem, any agent can match with any other agent. In a two-sided matching game, also known as the college admissions problem, agents are ex-ante placed into groups, such as applicants and colleges, and matches can only occur across, rather than within, groups. Given the lack of such ex-ante groupings, jurisdictional merger decisions are thus more appropriately modeled as one-sided. Fox (2005), Sorensen (2007), and Boyd et al (2003) provide econometric estimators for two-sided matching games. See Roth and Sotomayor (1990) for additional discussion of both one-sided and two-sided matching games.

⁷ That is, if all districts prefer any merger over remaining alone, 1 merging with 2 and 3 with 4 is a stable matching so long as there is no profitable deviation involving a merger between districts 2 and 4 (that is, $1 >_2 4$ or $3 >_2 4$). However, 1 merging with 4 and 2 with 3 is also a stable matching so long as there is no profitable deviation involving a merger between districts 1 and 3 (that is, $4 >_1 3$ or $2 >_3 1$).

⁸ In implementing this estimator, we incorporate geographic constraints by calculating utility gains to mergers only for adjacent pairs of districts. Each district then chooses partners from its geographically-constrained set of options. This is equivalent to setting $Q_{ji} = -\infty$ (or $U_{ji} = -\infty$) for non-adjacent districts.

cycling problem, we next introduce the restriction of *symmetry in match quality*:

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

That is, conditional on the attractiveness of a district, which is equally valued by each potential partner, and the inclination of a district to merge with any of its partners, the quality of the match is equally valued by the two districts i and j ; we again defer a discussion of the role of this assumption to the end of this section.⁹ Using this restriction, we have established the following result:

Proposition. *Under the assumptions of 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 regularity on preferences over merger partners in order to rule out cycling, which is the underlying source of the problems of non-existence and multiplicity.¹⁰ To see this, consider again the three district cycle 123 described above; this cycle can be equivalently represented by the following inequalities:

$$\begin{aligned} U_{21} &> U_{31} \\ U_{32} &> U_{12} \\ U_{13} &> U_{23}. \end{aligned}$$

Inserting our utility specification and imposing symmetry in match quality, we have that:

$$\begin{aligned} A_2 + Q_{12} &> A_3 + Q_{13} \\ A_3 + Q_{23} &> A_1 + Q_{12} \\ A_1 + Q_{13} &> A_2 + Q_{23}. \end{aligned}$$

Summing across these three inequalities, we thus have a contradiction ($A_1 + A_2 + A_3 + Q_{12} + Q_{23} + Q_{13} > A_1 + A_2 + A_3 + Q_{12} + Q_{23} + Q_{13}$), and it should be clear that cycling cannot occur under our assumption of symmetry in match quality.

While these existence and uniqueness properties are interesting from a theoretical perspective and, as noted above, are clearly important from an empirical perspective, the proposition is incomplete as we also need to characterize this unique stable matching in order to complete the development of our econometric estimator. Fortunately, under the assumptions of the proposition, symmetry in match quality and strict preferences, this unique stable matching can always be computed using the following simple algorithm:

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

Step B: Remove matched districts from map.

Step C: Drop matched districts from preference ranking and return to Step A.

⁹ This result is similar in flavor to that in Sorensen (2007), who shows that the condition of aligned preferences, which is equivalent to symmetry in utility, leads to a unique stable matching in two-sided matching games. He then develops a Bayesian estimator and applies this to the market for venture capital.

¹⁰ 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. To see this, note that if $U_{ij} = A_j + I_i + Q_{ij}$, where $Q_{ij} = Q_{ji}$, then preferences can be represented equivalently by $V_{ji} = U_{ji} - I_i + A_i$ and thus $V_{ij} = V_{ji}$. Under this alternative formulation, however, the gains to remaining unmerged ($V_{ii} = A_i - I_i$) does not necessarily equal zero. Said differently, we have normalized the utility from remaining unmerged to zero, whereas this alternative restriction of symmetric utilities allows for the utility to remaining unmerged to be arbitrary. Notably, however, neither restriction requires that the overall utility gains, relative to remaining unmerged, (i.e. $U_{ji} - U_{ii}$ or $V_{ji} - V_{ii}$) be symmetric between the two partners.

That is, each district ranks its potential merger partners according to its gain in utility from merging with each potential partner, and mutual first choices (including remaining unmerged) are matched with one another. After removing these matched districts from the map, mutual first choices from the reduced set of alternatives are matched, and the process continues until all districts are either matched with another district or remain unmerged. Again, the restriction of symmetry in match quality rules out cycles and thus guarantees that at least one border with two districts that are mutual first choices can be found in Step A. Without our symmetry assumption, the algorithm does not necessarily converge. For example, with the three-district cycle described above, no two districts are each other's first choice and the algorithm gets stuck at the first step. Our ability to calculate the stable matching via this simple iterative algorithm suggests that a simulation approach may be productive from an econometric perspective. After describing the role of the assumptions underlying these theoretical results, we next turn to the development of such an empirical approach.

3.2. Role of the assumptions

Several assumptions were required in order to generate these results of existence and uniqueness. In this section, we discuss the role of the key assumptions, both explicit, such as symmetry in match quality, and implicit, such as no side payments. Regarding the merger protocol, the key assumptions were no mergers involving more than two districts and no separations. Allowing mergers involving more than two districts would certainly complicate the problem as it would require consideration of many more combinations of merger partners and we would also need to generalize the payoff structure, which currently includes only own-district characteristics and those of a single neighbor. Given these issues and the difficulty of empirical implementation with only two-district mergers, we leave the issue of multiple merger partners for future work. Separations would also significantly complicate the analysis as it would require the researcher to identify borders within districts along which such separations may occur. Data from our empirical application to Iowa school districts generally support these assumptions, with only one case of a merger involving more than two districts in a given year and only two separations.

In our model, side payments were not allowed, and thus mergers that might increase overall surplus may be blocked by one of the two potential partners. In models with side payments, by contrast, the mergers that generate the largest joint surplus are the most likely to occur.¹¹ While introducing such side payments would clearly alter mergers that occur in a stable matching, we feel that this assumption is reasonable in our empirical application. Furthermore, any promises of side payments may suffer from credibility and enforcement problems in practice (Alesina and Spolaore, 2003).

Another implicit assumption is myopic decision-making: that is, we do not allow districts to consider how a merger today might alter the pool of potential merger partners in the future. While these dynamic considerations are certainly interesting, they would significantly complicate the analysis, and we thus leave them for future work. We do, however, update borders following mergers and allow these new districts to subsequently re-merge; such behavior is atypical in our sample, which contains only four such subsequent mergers.

Perhaps the most crucial assumption is symmetry in the match-specific quality component of the overall utility gain to a district from a merger. Even under the assumption of symmetry in match quality, the researcher can estimate a variety of econometric specifications controlling for own-district characteristics, characteristics of the other district, interactions of these characteristics across the two districts, and symmetric differences, such as squared or absolute differences, between the characteristics of the two districts. The assumption of

¹¹ In a model with side payments, Fox (2005) develops a non-parametric estimator for two-sided matching games using these efficiency properties of stable matchings.

symmetric match quality allows for asymmetries in the gains to a merger so long as these asymmetries are captured in the attractiveness and inclination values. Suppose, for example, that the utility gain to district i from merging with district j is increasing in $\beta(y_j - y_i)$, where y_i and y_j represent income in districts i and j , respectively and β is a parameter to be estimated. This formulation is fully consistent with our assumption of symmetry in match quality given that βy_j can be incorporated into attractiveness and $-\beta y_i$ can be incorporated into inclination.

4. Econometric implementation

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 and Z_i represents observed measures of district i 's inclination to merge with any of its potential partners, relative to remaining unmerged. 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. The vector $\theta = (\theta_x, \theta_z, \theta_w)$ represents parameters to be estimated. Symmetry in match quality is satisfied whenever $f(W_i, W_j) = f(W_j, W_i)$ and $\varepsilon_{ji} = \varepsilon_{ij}$; 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, no closed form solution exists for the probability of a merger between any two districts. Said differently, the probability of a merger between any two districts depends upon the characteristics of all districts, even non-adjacent ones. 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 first developed by Lerman and Manski (1981). In particular, for replication $r = 1, 2, \dots, R$, a symmetric unobserved match quality ($\varepsilon_{ji}^r = \varepsilon_{ij}^r$) 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 . This leads to the well-known frequency simulator:

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

where $y_{ij}^r \in \{0,1\}$ 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 $\Pr(y_{ij}^r = 1|\tau)$, where τ is the smoothing parameter, is preferred.¹² Importantly, as the smoothing parameter goes to zero, the smooth simulator approaches the frequency simulator [$\lim_{\tau \rightarrow 0} \Pr(y_{ij}^r = 1|\tau) = y_{ij}^r$]. But, for any positive value of the smoothing parameter, the smooth simulator is bounded

¹² A smoothed simulator is preferred to the frequency simulator for at least three reasons. First, mergers are relatively rare in practice, and thus, even with a large number of replications, borders may experience zero mergers across all replications and the frequency simulator will thus be zero. This creates problems for the GMM estimator, which places infinite weight on this hypothetical border observation given that the simulated variance is zero for this observation. Second, the theory underlying statistical inference for the GMM estimator assumes that the objective function is differentiable. Finally, a smooth objective function permits the use of computationally faster derivative-based optimization methods in choosing the parameters. See Stern (1997) for additional information on smooth simulators.

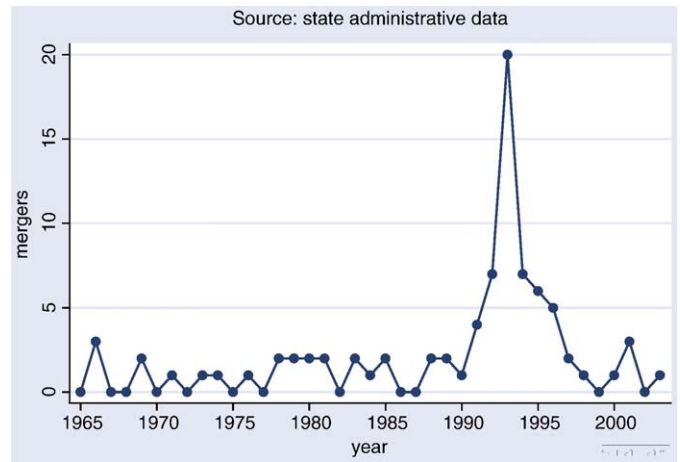


Fig. 2. Number of school district mergers in Iowa.

between zero and one ($\Pr(y_{ij}^r = 1|\tau) \in (0,1)$), and the average probability across all replications then serves as the estimate of the probability of a merger between i and j :

$$\widehat{Pr}(i, j) = \frac{1}{R} \sum_{r=1}^R \Pr(y_{ij}^r = 1|\tau)$$

We describe one possible smoothed simulator, which we use in the empirical application to follow, in Appendix B.

For estimation purposes, we use the method of simulated moments due to McFadden (1989).¹³ Under this method, parameters are chosen in order to minimize a measure of 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 moment conditions, and expressions for the variance-covariance matrix, are provided in Appendix C.

To summarize, estimation via simulation would proceed as follows:

- Step 0: For each border, independently draw an unobserved match quality (ε_{ij}) from the type-I extreme value distribution. Do this R times and index the replications $r = 1, 2, \dots, R$.
- 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.

The estimation approach is also summarized graphically via a flow chart in Fig. 3. As shown, the estimation involves both an inner loop, in which the simulated probabilities of merger are calculated given a set of parameters, and an outer loop, in which the set of parameters is chosen in order to minimize the GMM objective function.

An additional econometric issue raised by spatial interdependence involves the consistency of the estimator. Whether or not the estimator is consistent relates to the issue of whether the appropriate

¹³ Given the interdependence in merger decisions (if 1 merges with 2, then 3 cannot merge with 1 or 2) 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 is 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).

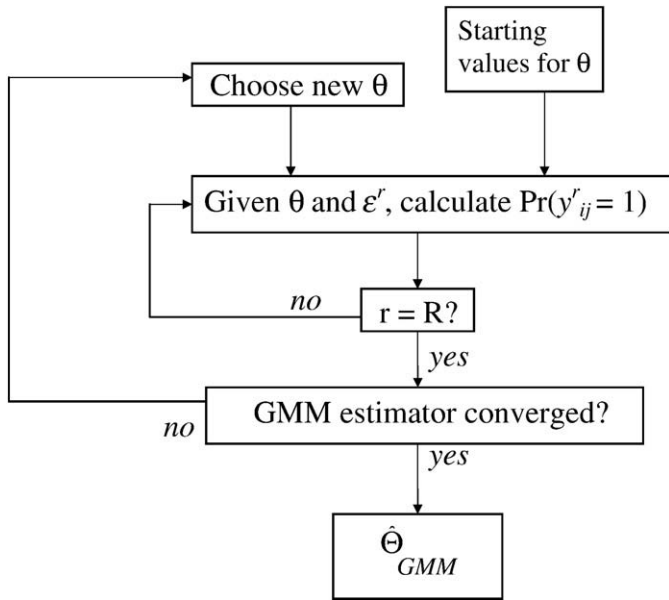


Fig. 3.

unit of observation is a single border on the comprehensive map. If the map is the unit of observation, then adding more borders does not necessarily add independent information, and the estimated parameters may not converge to the true parameters as the number of borders grows large. By contrast, we consider the border to be the appropriate unit of observation and, although there is certainly spatial dependence between borders, the estimator will be consistent under the condition of vanishing dependence, defined as the correlation between merger outcomes growing small as the distance between any two borders grows large.¹⁴ Moreover, vanishing dependence is likely to be satisfied under the requirement that mergers only occur between adjacent jurisdictions; this adjacency requirement is certainly satisfied in our empirical application, and we feel that it is typical of jurisdictional merger protocol. Simulations of a stylized version of our model under an adjacency requirement suggest that the condition of vanishing dependency is indeed satisfied: the correlation between mergers along any two borders is large with geographically close borders but grows small as the distance between any two borders grows large.¹⁵ Said differently, spatial interdependence appears to be a salient feature of the local, but not global, merger environment. This diminishing correlation should in turn induce the required degree of independence in border observations. While a formal exploration of this econometric issue is beyond the scope of

¹⁴ This condition of vanishing dependence is a central issue in the literature on spatial econometrics, which models the influence of behavior on neighbors. See, for example, Anselin (1988), Pinsky and Slade (1998), Pinsky et al. (2002), Kelejian and Prucha (2007).

¹⁵ In particular, we considered a linear map with 10 borders and 11 districts. District 1 can merge with 2 or remain unmerged, district 2 can merge with 1 or 3 or remain unmerged, and so on, up through district 11, which can merge with district 10 or remain unmerged. For each of 100,000 replications, we drew for each district a value for inclination and attractiveness from the standard normal distribution. For each border, we also drew a symmetric match quality from the standard normal distribution. Then, based upon these utility gains, we ran our iterative algorithm 100,000 times and computed the correlation between mergers for each border along the line. We found a correlation of roughly -0.26 for a distance of 1 border (e.g. the border between district 1 and 2 and the border between 2 and 3), reflecting the constraint that each district can only merge with one neighbor. The correlation is 0.06 for a distance of 2 borders. The positive correlation reflects the increased probability for a merger in the event of a merger two borders away. For example, a merger between districts 1 and 2 increases the probability of a merger between 3 and 4 given that 3 has lost its outside option of district 2. Beyond a distance of 2 borders, however, the correlation falls quickly to zero.

this paper, we feel that these simulation results should provide the reader with a degree of confidence in the consistency of our estimator.

Given these issues regarding the consistency of the estimator, we next examine the small sample properties of our estimator through a Monte Carlo analysis. Due to computational considerations, we keep the setup of this analysis simpler than in the general case, which, as noted above, has no closed-form solution for merger probabilities. Our estimator thus has both an inner loop, which calculates merger probabilities holding parameters fixed, and an outer loop, which chooses the parameter estimates that best match the mergers observed in the data. This estimation process requires several days to generate the main empirical results we present in this paper. A Monte Carlo analysis of this estimator would add a third loop, a set of replications. We thus provide a Monte Carlo analysis for a simple version of our model with a closed-form solution for the probability of a merger.

Consider three districts (1,2,3), each of whom can merge with any one of the others. In order to further simplify the model, we assume that each district would prefer any merger over remaining alone. Let the utility to i from a merger with j be given by:

$$U_{ji} = V_{ji} + \epsilon_{ji} = \beta_A A_j + \beta_Q Q_{ji} + \epsilon_{ji}$$

where $V_{ji} = \beta_A A_j + \beta_Q Q_{ji}$ and ϵ_{ji} represent the observed and unobserved utilities, respectively. Symmetry in match quality requires that $Q_{ji} = Q_{ij}$ and $\epsilon_{ji} = \epsilon_{ij}$. The unobserved utilities (ϵ_{ji}) are assumed to be distributed type-I extreme value. In this case, there are three possible outcomes: a merger between 1 and 2, a merger between 2 and 3, and a merger between 1 and 3. One district will always remain unmerged. In this case, the expression for the probability of a merger between any two districts has a closed-form solution. The probability of a merger between districts 1 and 2, for example, can be written as:

$$\begin{aligned} Pr(1,2) &= \frac{Pr(V_{21} + \epsilon_{21} > V_{31} + \epsilon_{31}, V_{12} + \epsilon_{12} > V_{32} + \epsilon_{32})}{1 + \exp(V_{31} - V_{21}) + \exp(V_{32} - V_{12})} \\ &= \frac{1}{1 + \exp[\beta_A(A_3 - A_2) + \beta_Q(Q_{13} - Q_{12})] + \exp[\beta_A(A_3 - A_1) + \beta_Q(Q_{23} - Q_{12})]} \end{aligned}$$

Thus, if $\beta_A > 0$ and $\beta_Q > 0$, the probability of a merger between districts 1 and 2 is declining in the attractiveness of 3, the quality of the match between 1 and 3, and the quality of the match between 2 and 3. The probability is increasing in the attractiveness of districts 1 and 2 and in the quality of the match between 1 and 2. While certainly simpler than the general model developed above, this special case retains the three key features that motivated our spatial merger estimator: two-sided decision making, multiple merger partners, and spatial interdependence. This latter feature is reflected in the inclusion of district 3 characteristics (A_3) in the probability of a merger between 1 and 2.

Table 1 provides results from our Monte Carlo analysis with true parameters $\beta_A = 0.2$ and $\beta_Q = 0.5$ and the number of markets equal to 15, 30, 60, or 120. Given the existence of a closed-form solution in this special case, we use a maximum-likelihood estimator. As shown, bias and RMSE for both the attractiveness and quality parameters are

Table 1
Results from Monte Carlo analysis.

# markets	Attractiveness parameter (β_A)		Quality parameter (β_Q)	
	Bias	RMSE	Bias	RMSE
15	.035	.426	.157	.480
30	-.012	.321	.069	.344
60	.001	.177	.022	.199
120	.013	.130	-.001	.122

Notes: The true parameters are $\beta_A = 0.2$ and $\beta_Q = 0.5$. We draw the observed values Q and A from the standard normal distribution once for each market. We draw the unobservable components of the model 200 times, once for each simulation. Our bias measure for β_A , our estimator for β_A , is given by $E(b_A - \beta_A)$, with a parallel expression for our quality parameter. The RMSE term for attractiveness is given by $[E(b_A - \beta_A)^2]^{1/2}$ with a parallel expression for quality. The parameters are estimated using maximum likelihood.

declining in the number of markets. Even with only 15 markets, however, the bias is relatively small, representing less than 20% of the value of the attractiveness parameter and about 30% of the value of the quality parameter. Moreover, moving from 15 to just 30 markets represents a further substantial improvement in the estimator. While our empirical application has only one geographic market, the state, we do have eleven time periods, which we treat as independent markets; we also have a much larger number of borders, which provides an additional source of variation. Of course, other applications of our spatial merger estimator may have significantly more geographic markets.

In summary, we have developed an econometric model of political integration that overcomes the three key limitations of existing econometric models. In particular, by appealing to the economics of matching and the associated stability concept, this 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 spatial interdependence of merger decisions. To illustrate the practical value of this approach, we next turn to an empirical application of school district mergers in the state of Iowa during the 1990s.

5. Empirical application

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. The number of school districts in the United States plummeted from around 120,000 in 1940 to under 15,000 today.

We choose to look at the experience of school districts in the state of Iowa during the 1990s for several reasons. Most importantly, while the state did provide financial incentives and technical assistance for consolidation, the decision to integrate ultimately rested with the school districts themselves. That is, school districts decided both whether or not to consolidate (in contrast to some other state mandates for minimum district size), and if they did choose to consolidate, they chose with which of their neighbors to do so (conditional on the neighbor's agreement).¹⁶ 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. Data on such boundaries would be difficult to assemble for earlier in the century.

Moreover, the consolidation environment in Iowa arguably satisfies two key assumptions underlying our methodology: symmetry in match quality and no side payments. We first discuss the symmetry assumption.

From our conversations with Guy Ghan, who oversaw the consolidation process for many of the districts in our sample, we understand that larger and more property-wealthy districts were considered to be universally attractive, rather than attractive in any match-specific way. In principle, one could include a measure of the income or wealth of the neighboring district in attractiveness and a measure of own income or wealth level into inclination. In practice, however, we have had difficulty estimating such asymmetries. They make the higher income district less likely to want to merge but the lower income district more inclined to merge, thus the overall effect on merger probabilities is ambiguous. We have had difficulty getting the estimator to converge when including such effects and therefore have followed the existing literature and included simpler heterogeneity measures. In order to estimate such effects, one could presumably use voting data from the referenda, which would separately reveal the preferences of the two districts. That is, one would expect to see more support in the lower income district and less support in the higher income district. Unfortunately, we have been unable to find such systematic voting data from Iowa.¹⁷

Regarding side payments, we know of no explicit monetary transfers between merging districts. Two possible alternative forms of side payments regard existing debt and the closing of schools. Regarding debt, two initial districts could continue to maintain separate tax rates in order to pay off the debt they had before merging. Another potential side payment involved discretion over which schools would be closed (particularly politically important for high schools) following a merger.¹⁸

The consolidation experience of Iowa school districts in the 1990s, while taking place in a less active period of consolidation nationally, shares key attributes of the environments in which the bulk of consolidation in the US has taken place. Changes in the agricultural industry led to shrinking school district enrollments. Economies of scale provided greater persuasive motivation for consolidation to state policymakers than to local voters who valued district identity and autonomy. The fact that local activity picked up after the state offered incentives is highly representative of the national experience. The incentives are also useful for our methodology, as they allow us to identify differential benefits to specific mergers. We next describe the data before turning to exact measures and the empirical results.

5.1. 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.¹⁹ 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. In order to use the 1989 Census data in analyzing mergers in each year of our panel, we created enrollment-weighted averages of those 1989 values for district-year observations that have experienced a merger between 1989 and the current year. These Census data include richer variables than found in the Common Core, including the distribution of adult educational attainment, self-reported home values, and area of the school district in school miles. We use the Common Core of Data for annual data on the number of total students enrolled in public school and

¹⁷ See *Filer and Kenny (1980)* for an analysis of the role of income differences in city-county consolidation referenda during the 1949–1976 period.

¹⁸ High school closure was not an issue for the majority of mergers in our sample, who shared one high school between the two districts prior to merging. Of the ten mergers with more than one high school between the pair going into the merger, all ten began with two high schools and eight of the ten mergers resulted in the closing of one of them.

¹⁹ School districts in Iowa are independent jurisdictions, meaning that they collect their own tax revenue rather than receiving revenue allocated to them by a parent government such as a town, city, or county.

¹⁶ The consolidation process in Iowa began with professional consultants working as contractors for the state drawing up reorganization feasibility studies and plans. Not every potential boundary was the subject of study, and not every studied merger went forward to the voting stage. We have limited data on reorganization referendum voting outcomes, but given the endogenous nature of even having a vote, we view actual consolidation, rather than vote shares, as the relevant outcome for our empirical analysis.

Table 2
Summary of merger incentives.

	WGS pre-91	WGS post-91	Reorganize pre-91 or post-93	Reorganize 91–93 (for 5 years after reorganization)
Foundation tax rate (paid to state on district assessed property valuation)	= \$5.40 per \$1000	= \$5.40 per \$1000	= \$5.40 per \$1000	Year 1 = 4.4 Year 2 = 4.6 Year 3 = 4.8 Year 4 = 5.0 Year 5 = 5.2
Per-pupil foundation payment from state to district (F)	F per non-WGS student 1.1F per WGS student	F per non-WGS student F per WGS student	F per non-WGS student F per previous WGS students	F per non-WGS student 1.1F per previous WGS student

enrollment by grade.²⁰ Data on school district finances are taken from the School District Finance Data (F33) file, available annually in our time period from the fall of 1989 to the fall of 2001. We use current instructional spending, 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. Relatedly, in two cases, individual districts were involuntarily dissolved into surrounding districts. Because these cases were both infrequent and involuntarily, we disregard them in our estimation, which allows a given district to remain unmerged or to merge with any one of its neighbors.

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.²¹ 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.

As mentioned above, our theoretical and econometric framework is purely static in nature: we do not allow districts to consider how a merger today might alter the pool of potential merger partners in the future. We stack our panel data and consider each year's map as a separate equilibrium of the merger game. In the construction of our dataset, we do, however, incorporate changes in potential merger partners following a merger. In particular, if two districts 1 and 2 merge in year *t* to form a new district 12, this new district 12 now shares borders with all of 1's original borders and 2's original borders, and we allow for such subsequent mergers between 12 and any of these potential merger partners in year *t* + 1.

5.2. Financial incentives

The state offered financial incentives to school districts voting by November 30, 1990 to make their consolidations effective between

²⁰ While race and ethnicity variables are available in the Census and the Common Core of Data, they provide limited variation in this instance and we exclude them from our analysis due to computational constraints.

²¹ School district boundaries in Iowa bear little relationship to county boundaries; that is, existing district boundaries often cross county lines, and mergers also occur across county lines.

Table 3a
Merger activity by district enrollments.

	Larger district < 600	600 <= larger district < 1200
600 <= smaller district < 1200		0.00%
Smaller district < 600	2.36%	0.36%

July 1, 1991 and July 1, 1993.²² As Fig. 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 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. 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 2. 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, taxpayers in districts with pre-merger enrollments of fewer than 600 students experienced a foundation tax rate of 4.40 mills in the first year post-consolidation, increasing by 0.20 mills per year until reaching 5.40 mills 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. The enrollment limit is defined separately for each of the two potential merger partners. All property in the post-merger district will be eligible for the lower foundation rate if each partner 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 and thus property owners in the two districts effectively pay different tax rates for the 5 years following a merger. We therefore measure this merger incentive separately for the two potential merger partners.

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 annual administrative data on assessed property values. We then compute the present discounted value of the five-year stream of payments using an assumed discount rate of 3%, which is roughly the inflation rate during 1991, and, given the stagnant population in Iowa, an assumed nominal growth rate in housing values of zero.

As shown in Table 3a, 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. While these average merger rates of 2.4% appear relatively low at first glance, it is important to note that these merger rates are both along a given border and within a given year. Districts with enrollments below

²² Numerous state publications and media coverage suggest that the state legislature promoted consolidation because members believed it would improve school quality, rather than that it would significantly reduce per-pupil costs. The state department of education publishes an Annual Condition of Education Report which reports the distribution of curricular programs offered, achievement levels, and fiscal status by district size. This report presents descriptive tables and figures, without testing for statistical differences across enrollment categories. Smaller districts are shown to have more limited high school subject offerings, lower ACT scores, and lower preschool enrollment rates. Given the already widespread use of whole grade sharing arrangements, in which two districts remain autonomous but educational services for a given grade level are provided to all students by only one district, it is still unclear why the state was willing to finance the push for fiscal consolidation.

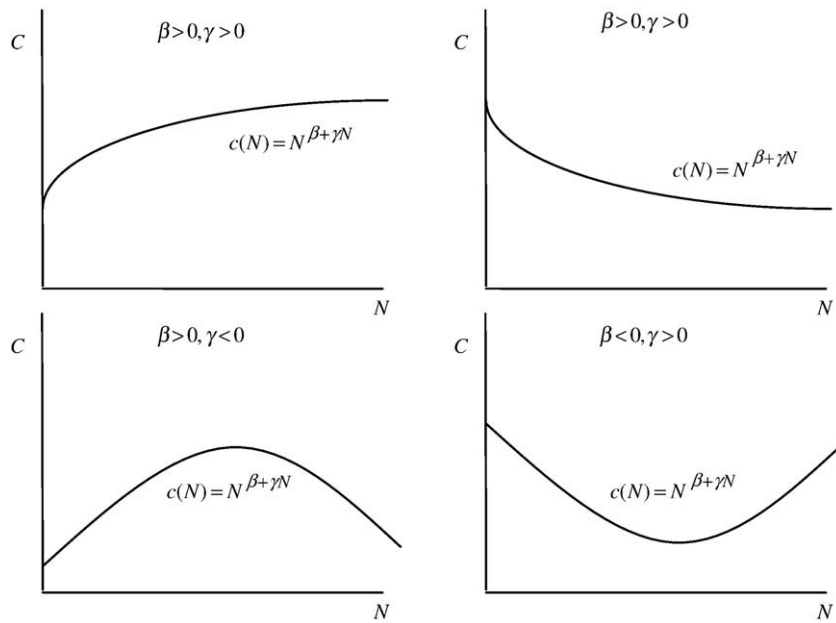


Fig. 4. Cost curve specification.

600 merged with one of their multiple neighbors at an 8% rate in a given year during the 1991–1993 period. Even this district-level rate understates merger activity given that no subsequent mergers occurred during the 1991–1993 period. That is, almost one-third (31%) of districts with enrollments below 600 in 1991 merged with one of their neighbors at some point during the 1991–1993 period. Taken together with the spike in mergers during this incentive period, as demonstrated in Fig. 2, this evidence suggests that districts strongly responded to the financial incentives in place during this period. In addition, as will be argued below, we can use the 600-cutoff level to distinguish the role of the fiscal incentives from the role of district enrollment.

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 schools, 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 5 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.²³ Unlike the

first incentive, which varied between the two potential partners, this second incentive provided extra funds directly to the new district, and both districts would thus share in the incentive post-merger. Many of the districts consolidating, both during the 1991–1993 eligibility window and afterwards, had been involved in WGS agreements.²⁴

5.3. Heterogeneity factors

We focus on three measures of heterogeneity: fiscal, demographic, and spatial. These latter two measures are emphasized in the work of Alesina and Spolaore (1997, 2003). As a baseline measure of fiscal heterogeneity, we use the squared 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 squared difference in the measures between the two adjacent districts. For our demographic heterogeneity measures, we examine the squared difference across the two districts in two measures of adult educational attainment: the percent of adults with less than a high school degree, and the percent of adults with at least a four-year college degree. Finally, regarding spatial heterogeneity, we control for the population density of the district, as measured in total population per square mile, as well as the estimated distance between the two districts.²⁵

²³ 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% 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% and a nominal growth rate in the foundation payment of 4.5%, which is roughly the growth rate realized during this period.

²⁴ 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.

²⁵ We estimate the distance between the two districts using data on each district's area in square miles, with the simplifying assumption that each district is square in shape. This allows us to calculate the distance between the two districts (from center to center) as the sum of one-half of the square root of each district's area.

Table 3b

Merger activity by district enrollments, further details.

	Larger district < 300	300 ≤ larger district < 600
300 ≤ smaller < 600		1.79%
Smaller district < 300	3.90%	2.49%

5.4. 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:

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

As shown in Fig. 4, this specification allows for a wide range of shapes for the cost curve. That is, the parameter β describes the shape of the cost curve at low enrollment levels, while the parameter γ captures the shape of the cost curve at high enrollment levels. Thus, the former parameter can be interpreted as a measure of economies of scale for small districts and the latter can be interpreted as a measure of potential diseconomies of scale for large districts. As shown, if $\beta < 0$ but $\gamma > 0$, the cost-curve will be U-shaped, suggesting that mergers will be efficiency enhancing for smaller districts but potentially efficiency detracting for larger districts. Inserting this cost curve specification into our measure of efficiency gains, we have that:

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

Thus, as described above, our estimate of β can be considered an estimate of the role of economies of scale for small districts in merger decisions, while our estimate of γ can be considered a corresponding estimate of the role of diseconomies of scale for large districts. Note that the measure of efficiency gains satisfies our assumption of symmetry in match quality given that the two own-district terms $\ln(N_i)$ and $N_i \ln(N_i)$ can be included in the inclination measure.

Given that the merger incentives are targeted at small districts, it is important to be clear in describing how we distinguish between responsiveness to the incentives and the role of economies of scale. First, we have time-series variation in the merger incentives, which were available only during the 1991–1993 period, and, as shown above, merger rates were much higher during this period. A second source of identification is the 600–cutoff level for the merger incentives. As shown in Table 3b, merger rates were much higher along borders with two very small districts, defined as those with enrollments below 300, suggesting a role for economies of scale. Roughly speaking, we can use the variation in merger rates within the below-600 category to identify the role of economies of scale, while variation in merger rates between the below-600 and above-600 categories identifies the responsiveness of districts to the merger incentives.

Table 4

Summary statistics for key variables.

Observation = school district / year	Mean (standard deviation)	
	Merger n = 108	No merger n = 4579
Enrollment	452 (311)	1288 (2499)
Percent of adults with less than HS degree	0.2091 (0.0459)	0.2062 (0.0529)
Percent of adults with college degree	0.1066 (0.0280)	0.1178 (0.0529)
Per-pupil instructional spending (scaled by housing values)	0.0801 (0.0295)	0.0703 (0.0288)
Area (square miles)	98.7517 (43.0814)	134.9075 (69.1751)
Population density	4.5581 (2.2789)	17.3308 (54.3544)

5.5. Results

Table 4 provides summary statistics for our key variables at the level of an individual school district conditional on whether or not a merger occurred. As shown, those districts choosing to merge with one of their neighbors were significantly smaller than those choosing to remain unmerged. Moreover, these districts tended to be significantly smaller in a geographic sense, as captured by square miles, suggesting an important role for spatial heterogeneity and transportation costs. As shown in the final row, however, merging districts tend to be less dense as the effect of low enrollment dominates the effect of small size in the geographic sense.

Similarly, Table 5 provides summary statistics at the level of a school district border. As shown, mergers were more likely to occur along borders that were eligible for the merger incentives, and this difference is statistically significant. 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 adult educational outcomes although this difference is only statistically significant for the college degree heterogeneity variable. Regarding the second measure of heterogeneity, differences in spending levels, no effect is detected in these summary statistics. Finally, mergers were more likely to occur among geographically close

Table 5

Summary statistics for key variables.

Observation = border/year	Mean (standard deviation)	
	Merger n = 50	No merger n = 12184
Merger incentive	2.4172 (2.4516)	0.4737 (1.1630)
Economies of scale	− 0.7772 (0.4422)	− 0.8420 (0.6197)
Diseconomies of scale	− 3140.61 (2067.28)	− 11969.15 (28582.14)
Heterogeneity in percent adults with less than HS degree	0.0030 (0.0036)	0.0032 (0.0049)
Heterogeneity in percent adults with college degree	0.0015 (0.0020)	0.0037 (0.0098)
Heterogeneity in spending (property-value adjusted)	0.0010 (0.0028)	0.0009 (0.0033)
Estimated distance between districts	9.6427 (1.1579)	11.4768 (2.1751)
Population density	4.4837 (1.9605)	14.9497 (46.2279)

Table 6
Determinants of school district consolidations.

	Column 1	Column 2	Column 3	Column 4	Column 5
Merger incentive	0.5139** (0.1347)	0.5177** (0.1244)	0.5179** (0.1139)	0.3235** (0.1354)	0.5234** (0.0945)
Economies of scale	-0.3184** (0.0249)	-0.2992** (0.0211)	-0.3066** (0.0189)	-0.2106** (0.0347)	-1.0722** (0.0382)
Diseconomies of scale	0.0002** (0.0000)	0.0002** (0.0000)	0.0002** (0.0000)	0.0002** (0.0000)	0.0002** (0.0000)
Heterogeneity in percent adults with less than HS degree		-6.8629 (8.9018)			
Heterogeneity in percent adults with college degree		-11.0788** (2.2978)			
Heterogeneity in spending (property-value adjusted)			-68.6160** (10.4858)		
Estimated distance between districts				-0.5045** (0.0360)	
Population density				-0.1184** (0.0004)	
Constant	-6.7812** (0.0717)	-6.7925** (0.0654)	-6.7383** (0.0626)	-8.1039** (0.0748)	-6.7316** (0.0428)
Sample size	12,234	12,234	12,234	12,234	12,234
Serial correlation assumption	Independent	Independent	Independent	Independent	Dependent

Standard errors below coefficients.

*Significant at 10%; **significant at 5%.

districts, and this difference is statistically significant. We next turn to a more formal econometric test of our hypotheses.

Table 6 provides the results from our simulated method of moments estimator. Column 1 presents our baseline results, which do not include any heterogeneity measures, while columns 2–4 introduce our three different heterogeneity measures. As shown in the baseline results, merger incentives have a positive effect on the decision to merge, providing evidence that is consistent with the suggestive evidence provided in Fig. 2 and Table 3a. While these results demonstrate a positive and statistically significant effect of financial incentives on merger decisions, Table 7 provides some evidence on the magnitude of the effect. In particular, we conduct a counterfactual experiment in which the financial incentives in place during 1991–1993 are eliminated. As shown, merger rates along borders in which both initial districts had enrollments below 600 and were thus eligible for the foundation rate incentives fall significantly from 2.89% to 0.46%, representing an 84% reduction in these merger rates.

Returning to the coefficients in column 1 of Table 6, the economies of scale variable has the expected negative coefficient, while the diseconomies of scale measure has an expected positive coefficient. The magnitude of these coefficients, however, are unrealistic, implying that average costs are minimized at enrollments of about 250 students.²⁶ Estimates of education cost functions, as summarized by Andrews et al. (2002), imply that diseconomies of scale may not set in until enrollments reach 6000 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. In addition to transportation costs, two other factors likely contribute to our lower estimates. First, 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. Second, we generate these revealed preference estimates of optimal district size from a sample with smaller districts than the national population of districts. As shown in columns 2–4, the coefficients on the merger incentive measure and the size variables are similar after controlling for our various measures of heterogeneity. Turning to these measures, we find that all three sources of heterogeneity served as a

repelling force in merger decisions, and all of these coefficients are statistically significant except for the adult high school degree heterogeneity measure. In particular, mergers are less likely to occur as the difference in the fraction of adults with a college degree increases, as shown in column 2. Similarly, as the difference in property-value adjusted spending levels increases, the propensity to merge falls, as shown in column 3. Finally, as shown in column 4, estimated distance between the two districts, conditional on population density, also serves as a repelling force, presumably due to the higher transportation costs. To provide a sense of the magnitude of these effects, Table 7 provides results from simulations in which the heterogeneity channel is shut down. For example, as shown in experiment 3, merger rates increase from 0.43% to 0.46%, an increase of 7%, if districts do not condition on spending differences, or equivalently, if these disparities between districts are eliminated. Similarly, setting the distance between two districts equal to zero leads to an economically significant increase in merger rates.²⁷

One implicit econometric assumption underlying the results in columns 1–4 of Table 6 is a lack of serial correlation. That is, we have assumed that the unobserved match qualities are independent over time for a given border, whereas it might seem more reasonable to assume that these match qualities are correlated over time. In order to test the robustness of this assumption, column 5 of Table 6 presents results from an alternative specification in which unobserved match qualities are drawn for a given border in 1991 and are then held constant over the remaining sample period.²⁸ As shown, the coefficients and standard errors are qualitatively similar to those in our baseline results of column 1, suggesting that our assumption of independence is not driving the statistical significance of our baseline empirical results.

In order to compare the results of our estimator relative to those in the existing literature, Table 8 presents the results from a bivariate logit

²⁷ The large effect of estimated distance in experiment 4 may be due to two factors. First, this experiment holds constant population density, which would obviously tend to increase as the geographic size of the two districts is decreased. Second, there is likely multicollinearity in the underlying specification in Table 6 as it simultaneously controls for enrollment, distance between districts and population density.

²⁸ We treat new districts following mergers as having new borders. That is, if two districts 1 and 2 merge, creating district 12, then any borders of the new district 12 are treated as new borders and a new unobserved match quality is drawn from the type-I extreme value distribution.

²⁶ Recall that our assumed cost curve is given by $c(N) = N^\beta + \gamma N$.

Table 7
Counterfactual experiments.

	Merger rates	
	Baseline	Counterfactual
Experiment 1: remove merger incentives, 1991–1993 (merger rates for borders with both districts below 600 enrollment)	2.89%	0.46%
Experiment 2: eliminate adult educational heterogeneity	0.43%	0.43%
Experiment 3: eliminate heterogeneity in spending levels	0.43%	0.46%
Experiment 4: set distance to zero	0.41%	11.54%

model under the assumption of symmetry in match quality.²⁹ As shown in column 1, the merger incentives coefficient is positive and statistically significant. While this effect has the same sign as the corresponding coefficient in Table 6, it is difficult to compare the magnitude of the policy incentives effect in the two models given the independence assumption in the bivariate model. In particular, the results from policy experiments comparable to those in Table 7 would be difficult to interpret as there is no requirement that each district merge with only one partner in the bivariate probit model. While the size of these policy incentive effects are not directly comparable, the role of scale is more comparable: the bivariate logit coefficients suggest a minimum efficient size of 688, significantly larger than that suggested via our simulation approach. Regarding the role of heterogeneity, the bivariate probit finds statistically significant support for only one out of the three measures, and this does not appear to be driven by differences in power as the t-statistics on own-district characteristics, such as the merger incentive, are similar to those in Table 6. As shown in column 4, geographic distance between districts creates a disincentive to merge, while columns 2 and 3 report no role for heterogeneity in the educational attainment of parents or differences in spending levels. One interpretation of these weaker heterogeneity results is that our spatial merger estimator, but not the bivariate logit model, accounts for the fact that districts choose one district from their many potential merger partners and thus more naturally allows for districts to choose the one district with which it is most appropriately matched.

6. Conclusion

In this paper, we develop an empirical approach to the study of jurisdictional 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. Applying this method to a spate of school district mergers in Iowa during the 1990s, our results demonstrate the importance of state subsidies and the limited role of heterogeneity in explaining the patterns of mergers in Iowa during this time period. In a comparison with existing estimators, we demonstrate the relative power of our estimator to detect a statistically significant role for heterogeneity factors. One caveat is that the results of this

²⁹ There are two differences between our bivariate logit model and the bivariate probit model estimated by Brasington. First, while Brasington’s model estimates a parameter capturing the correlation between the unobserved utility of the two districts, we simply assume that this correlation equals 1 (we attempted to estimate a model with such a correlation but had difficulties with convergence as the correlation tended towards the extreme values of +1 or –1). Second, we use the logistic distribution rather than the normal distribution in order to get a closed form expression for the merger probability and to make our results more comparable with our simulation estimator, which assumes that match qualities are distributed type-I extreme value. As is well known, however, the normal and logistic distributions are both symmetric and are quite similar in shape, suggesting that our results are probably invariant to this distributional assumption.

Table 8
Results from bivariate logit model.

	Column 1	Column 2	Column 3	Column 4
Merger incentive	0.4226** (0.0762)	0.4219** (0.0762)	0.4146** (0.0754)	0.2982** (0.0792)
Economies of scale	–2.5942** (0.9728)	–2.5964** (0.9735)	–2.4376** (0.9090)	–1.0391 (1.5017)
Diseconomies of scale	0.0005** (0.0001)	0.0005** (0.0001)	0.0005** (0.0001)	0.0001 (0.0003)
Heterogeneity in percent adults with less than HS degree		–7.9141 (37.0851)		
Heterogeneity in percent adults with college degree		–1.9801 (60.5672)		
Heterogeneity in spending (property-value adjusted)			–117.6041 (78.9324)	
Estimated distance between districts				–0.5644** (0.2149)
Population density				–0.1560 (0.1425)
Constant	–4.9715** 0.4492	–4.9401 (0.4719)	–4.6454** (0.4596)	1.5591 (2.8054)
Sample size	12,234	12,234	12,234	12,234

Standard errors below coefficients.
*Significant at 10%; ** significant at 5%.

analysis, which abstracts from racial heterogeneity, may not generalize to other states and time periods. Iowa has very little racial heterogeneity, and, as noted above, other studies, such as Alesina et al. (2004), have found a strong role for such heterogeneity in terms of predicting the number of school districts within U.S. counties. The methodology we develop, however, could be adapted to consider racial and other sources of heterogeneity. We anticipate that this methodology will be useful in analyzing spatial integration in a variety of applications, particularly those analyzing more recent and ongoing boundary changes, which benefit from the availability of geocoded boundary data.

Appendix A

Proof of Proposition. Chung (2000) and Tan (1991) have shown that the absence of odd cycles, as defined in the text, implies the 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 proof shows that under the restriction of symmetry in match quality and strict preferences, there are no cycles. Thus, under 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 ($2 \succ_1 4$). In order for matching B to be stable, it must be the case that district 2 is paired with another district, which we denote district 3, and further that district 2 prefers 3 over 1 ($3 \succ_2 1$).³⁰ 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 merges with district 5 in matching A , it must be the case that 3

³⁰ If district 2 is paired with itself in matching B , it must prefer district 1, its partner in matching A ($1 \succ_2 2$) in order for matching A to be stable. But it is then clear that matching B is unstable as district 1 prefers 2 over 4.

prefers 5 over 2 ($5 >_3 2$) 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 ($6 >_5 3$). 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 ($4 >_6 5$) in order for A to be stable. But, in order for B to be stable, 4 must prefer 1 over 6 ($1 >_4 6$) 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:

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

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

$$\begin{aligned} A_2 + Q_{1,2} &> A_C + Q_{1,C} \\ A_3 + Q_{2,3} &> A_1 + Q_{1,2} \\ A_4 + Q_{3,4} &> A_2 + Q_{2,3} \\ &\dots \\ A_C + Q_{C,C-1} &> A_{C-2} + Q_{C-1,C-2} \\ A_1 + Q_{1,C} &> A_{C-1} + Q_{C,C-1}. \end{aligned}$$

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 B. 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_j^* and U_i^* as the equilibrium utility, or value of the game, for districts i and j under the stable matching. After calculating these utilities, we provide each district a small amount of additional information ($\tau\eta_{ij}$), which is also distributed type-I extreme value and is assumed symmetric between the two partners, regarding each of their merger 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:

$$\begin{aligned} Pr(\text{deviation}_{ij}) &= Pr(U_{ij} + \tau\eta_{ij} > U_j^* + \tau\eta_j, U_{ji} + \tau\eta_{ij} > U_i^* + \tau\eta_i) \\ &= \frac{1}{1 + \exp\left[\frac{(U_j^* - U_{ij})}{\tau}\right] + \exp\left[\frac{(U_i^* - U_{ji})}{\tau}\right]} \end{aligned}$$

where τ is the smoothing parameter and is chosen to be small. As this smoothing parameter converges to zero, the probability of a merger approaches 0 given that at least one district must prefer its equilibrium utility to this deviation option in order for this matching to be stable ($U_j^* > U_{ij}$ or $U_i^* > U_{ji}$). 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 from the stable matching, where the probability of a deviation is increasing in the smoothing parameter τ and decreasing in the difference between equilibrium values and deviation values ($U_j^* - U_{ij}$ and $U_i^* - U_{ji}$). For merged borders ($y_{ij} = 1$), create a set of willing deviation partners for i ($U_{ik} > U_k^*$); to guarantee that this set is not empty, include district i itself in this set where $U_{ii} = 0$. Denote this set as B_i with elements indexed by k . Create a

similar set for j ; denote this set B_j with elements indexed by l . Then, we have that the merger probability can be written as follows:

$$\begin{aligned} Pr(\text{no deviation}_{ij}) &= \\ &Pr(U_{ji} + \tau\eta_{ij} > U_{ki} + \tau\eta_{ki} \text{ all } k, U_{ij} + \tau\eta_{ij} > U_{lj} + \tau\eta_{lj} \text{ all } l) \\ &= \frac{1}{1 + \sum_k \exp\left[\frac{(U_{ki} - U_{ji})}{\tau}\right] + \sum_l \exp\left[\frac{(U_{lj} - U_{ij})}{\tau}\right]}. \end{aligned}$$

Again, as the smoothing parameter converges to zero, the probability of no deviation, and thus a merger, approaches 1 as each term in the summation (for example, $\exp[(U_{ki} - U_{ji})/\tau]$) approaches zero given that district i must prefer j over k in order for the matching to be stable. Thus, the smooth simulator again approaches the frequency simulator.

Appendix C. 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. N thus represents the number of borders across both time and space. Finally, W is a $k \times k$ weighting matrix. The optimal weighting matrix is given by the inverse of the variance-covariance matrix:

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

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:

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

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:

$$\text{Var}(\theta) = (1/N) I' [dp/d\theta] W [dp/d\theta] I$$

References

Alesina, Alberto, Spolaore, Enrico, 2003. *The Size of Nations*. Cambridge, MIT Press.
 Alesina, Alberto, Spolaore, Enrico, 1997. On the size and number of nations. *Quarterly Journal of Economics* 112, 1027–1056.
 Alesina, Alberto, Baqir, Reza, Hoxby, Caroline, 2004. Political jurisdictions in heterogeneous communities. *Journal of Political Economy* 112, 348–396.
 Andrews, Matthew, Duncombe, William, Yinger, John, 2002. Revisiting economies of size in American education: are we any closer to a consensus? *Economics of Education Review* 21, 245–262.
 Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Boston, MA.
 Bolton, Patrick, Roland, Gerald, 1997. The breakup of nations: a political economy analysis. *Quarterly Journal of Economics* 112, 1057–1090.
 Boyd, Donald, Lankford, Hamilton, Loeb, Susanna, Wyckoff, James, 2003. Analyzing the determinants of the matching public school teachers to jobs: estimating compensating differentials in imperfect labor markets. NBER working paper, vol. 9878.
 Brasington, David, 1999. Joint provision of public goods: the consolidation of school districts. *Journal of Public Economics* 73, 373–393.
 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.
 Brasington, David, 2003b. Size and school district consolidation: do opposites attract? *Economica* 70, 673–690.
 Chung, Kim Sau, 2000. On the existence of stable roommate matchings. *Games and Economic Behavior* 33 (2), 206–230.

- Dranove, David, White, William, 1994. Recent theory and evidence on competition in hospital markets. *Journal of Economics and Management Strategy* 3 (1), 169–209.
- Filer, John., Kenny, Lawrence, 1980. Voter reaction to city–county consolidation referenda. *Journal of Law and Economics* 23 (1), 179–190.
- Fox, Jeremy, 2005. Nonparametric estimation of matching games with endogenous prices. working paper. University of Chicago.
- Gordon Nora, Knight Brian, 2008. The effects of school district consolidation on educational cost and quality. *Public Finance Review* 36 (4), 408–430.
- Hooker, Clifford, Mueller, Van, 1970. The relationship of school district reorganization to state aid distribution systems. National Education Finance Project Special Study No. 11.
- Kelejjan, Harry, Prucha, Ingmar, 2007. HAC estimation in a spatial framework. *Journal of Econometrics* 140 (1), 131–154.
- Kenny, Larry, Schmidt, Amy, 1994. The decline in the number of school districts in the U.S.: 1950–1980. *Public Choice* 79, 1–18.
- Lerman, R., Manski, C., 1981. On the use of simulated frequencies to approximate choice probabilities. In: Manski, C., McFadden, D. (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge.
- McFadden, Daniel, 1989. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* 57, 995–1026.
- Nelson, Michael A., 1990. Decentralization of the subnational public sector: an empirical analysis of the determinants of local government structure in metropolitan areas in the U. S. *Southern Economic Journal* 57 (2), 443–457.
- Persson, Torsten, Tabellini, Guido, 2000. *Political Economics: Explaining Economic Policy*. MIT Press, Cambridge.
- Pinske, Joris, Slade, Margaret, 1998. Contracting in space: an application of spatial statistics to discrete choice models. *Journal of Econometrics* 85, 125–154.
- Pinske, Joris, Slade, Margaret, Brett, Craig, 2002. Spatial price competition: a semiparametric approach. *Econometrica* 70 (3), 1111–1153.
- Poirer, Dale, 1980. Partial observability in bivariate probit models. *Journal of Econometrics* 12, 209–217.
- Rodrigues-Neto, Jose, 2007. Representing roommates preferences' with symmetric utilities. *Journal of Economic Theory* 135, 545–550.
- Roth, Alvin, Sotomayor, Marilda A. Oliveira, 1990. *Two-sided matching: a study in game-theoretic modeling and analysis*. Cambridge University Press, Cambridge.
- Sorensen, Morten, 2007. How smart is smart money: a two-sided matching model of venture capital. *Journal of Finance* 62, 2725–2762.
- Stern, Steven, 1997. Simulation-assisted estimation. *Journal of Economic Literature* 35 (4), 2006–2039.
- Tan, Jimmy J.M., 1991. A necessary and sufficient condition for the existence of a complete stable matching. *Journal of Algorithms* 12 (1), 154–178.