Estimating a Game of Managing School District Capacity as Parents Vote with Their Feet*

Dennis Epple  
Carnegie Mellon University  
Akshaya Jha  
Stanford University  
Holger Sieg  
University of Pennsylvania  

August 25, 2014  

*We would like to thank Lanier Benkard, Eric Bettinger, Flavio Cunha, John Engberg, Kirill Evdokimov, David Figlio, Eric Hanushek, Dave Marcotte, Francois Margot, Mark Rosenzweig, Petra Todd, Frank Wolak, and seminar participants at various universities and conferences for comments. Financial support for this research is provided by the Institute of Education Sciences (IES R305A070117).
Abstract

Many urban districts confront the necessity of closing schools due to declining enrollments. To address this important policy issue, we formulate a sequential game of managing a school district’s student capacity. We show that a Perfect Bayesian Equilibrium exists and characterize its properties. Using data for a mid-sized district with declining enrollments, we estimate the parameters of our model. We show that consideration of student sorting is vital to the assessment of any school closing policy. We find that the district can reduce excess school capacity by closing underperforming schools. This approach inevitably leads to disruptive displacement of students in closed schools. Moreover, a significant fraction of students leaves the district. We also show that superintendents confront a difficult dilemma: pursuing an equity objective, such as limiting demographic stratification across schools, results in the exit of many more students than are lost by an objective such as closing underperforming schools.
1 Introduction

Student retention has increasingly become an important issue for urban school districts in the U.S., especially in the East and Midwest. When policies aimed at retaining students are not successful, urban school districts are forced to downsize. Between 2001 and 2009, Chicago closed 44 schools. The district recently announced closing of an additional 49 schools. Detroit closed more than 100 schools over the last decade. Philadelphia recently announced closure of 23 schools. Kansas City, Milwaukee, Pittsburgh, and Washington closed between 20 and 30 schools each in recent years.\(^1\) There are important research questions that arise in the context of downsizing a school district that have not been studied in the previous literature.\(^2\)

We develop a game in extensive form that captures the interactions between a superintendent, who manages the capacity of a school district, and parents, who choose schools for their children. We estimate the model using data from one school district that closed a significant number of schools to reduce excess capacity. The district sought to close underperforming schools. We quantify the effects of closing schools using the school district’s quality measures and objective function.\(^3\) We then solve the model using alternative objectives for the district. These counterfactual policy experiments illustrate the key trade-offs faced by the superintendent. While a district can reduce excess school capacity by closing underperforming schools, this approach inevitably leads to displacement of students in closed schools. Moreover, school closings induce some students to leave a district. We also show that superintendents confront a difficult dilemma: pursuing an

---

\(^1\)These cases are discussed in Dowdall (2011), Hurdle (2013) and Ahmed-Ullah, Chase, and Secter (2013).

\(^2\)There is a small empirical literature that has focused on quantifying the impact of school closing on student achievement. See, for example, Engberg, Gill, Zamora, and Zimmer (2012).

\(^3\)“Underperforming” schools were identified based on school fixed effects in student-level test score regressions.
equity objective, such as limiting demographic stratification across schools, results in the exit of many more students than are lost by objectives that are more attuned to parental preferences. More promising outcomes can be obtained by maximizing retention of students in the district or minimizing the number of students in schools chosen for closure. Summarizing, we provide and implement a new quantitative framework for evaluating the impact and effectiveness of different downsizing policies.

Our analysis is based on a new game in extensive form that captures the key problems associated with managing school district capacity. In the first stage of our game, the superintendent chooses schools to close to achieve a capacity-reduction requirement. In the second stage, parents choose among public schools in the district. Parents also have outside options: they can send their children to a private school within the district or leave the district altogether. Each public school is treated as a differentiated product that can be characterized by a vector of endogenous peer measures and other exogenous characteristics. We focus on Perfect Bayesian Equilibria (PBE).

We show that an equilibrium of the game can be characterized using a standard backward induction argument. Given a school configuration in the second stage, parents have beliefs about the peer qualities of each school and choose the optimal school for their children. Each parent has a vector of idiosyncratic shocks that are private information. While parents play pure Nash strategies, their behavior is random from the perspective of the superintendent. Equilibrium in the second stage subgame requires that parents’ beliefs about school quality are consistent with equilibrium sorting. We show that equilibria exist for this subgame. However, there is some scope for multiplicity.

The superintendent maximizes his or her objective function subject to the individual rationality and feasibility constraints implied by parental sorting in the second stage of the game, school-specific capacity constraints, and a constraint that bounds the admissible excess capacity in the system. We define a Perfect Bayesian Equilibrium for the game.
and show that it exists. However, non-uniqueness of equilibrium in the second stage potentially induces non-uniqueness of the PBE of the full game. To solve this problem, we propose a refinement concept that restricts attention to equilibria that are efficient from the perspective of the superintendent. Using this equilibrium selection rule, there exists a unique efficient PBE of the game.

The equilibrium of this model cannot be cast as the solution to a simple assignment problem. The self-selection constraints that arise due to the equilibrium sorting of students in the second stage of the game depend on the control variables in a nonlinear way. Different school configurations imply different levels of school qualities. We can show that computing an equilibrium of our game is equivalent to solving a complicated integer programming problem with non-linear constraints. There is a curse of dimensionality encountered in computing equilibria for this model since the number of possible school closings increases very rapidly as we increase the desired reduction of capacity (so-called “combinatorial explosion”). With presently available methods, we cannot solve problems that involve closing a large number of schools. As we discuss later, however, development of computational methods for this class of problems is an active area of investigation in operations research. It is possible to compute fully solved equilibria under various objectives for moderate school capacity reductions and thereby highlight key issues facing superintendents in practice as they undertake school closings.

To implement our quantitative framework, we need to estimate the parameters of our model. Our approach to demand estimation follows standard practices in the literature on differentiated products. The parameters characterizing parents’ preferences can be
estimated using a two-step estimator that controls for omitted school characteristics.\textsuperscript{4} The first stage is panel data discrete choice estimator which allows for school fixed effects. Most parameters of the model are identified in this stage.

The second stage of the estimator decomposes the school specific mean utilities (fixed effects) into its components. The main difficulty encountered in this stage of the estimation is that most observed school characteristics are peer measures and, therefore, endogenous outcomes of the sorting process. We apply an instrumental variable estimator to account for the endogeneity of school characteristics. Instruments are based on peer effects predicted by administrative school assignment rules that are implemented by the school district to predict the composition of each school before and after the school closing.\textsuperscript{5} The school assignment rules used by our district are primarily a function of distance, i.e. the district tends to assign a student to the school that is closest to the home of residence. These predicted compositions of schools can, therefore, be viewed as nonlinear transformation of historical, spatial residential sorting patterns.\textsuperscript{6}

Our instruments are valid under a number of plausible assumptions. We show in this paper that schools in the district experienced large changes prior and after the school closing. These changes were largely caused by a raptly shrinking district wide enrollment. During the period leading up to the school closing, the district bore 75 percent of the

\textsuperscript{4}The basic random coefficient model with unobserved product characteristics is introduced by Berry (1994) and Berry, Levinsohn, and Pakes (1995). Berry, Levinsohn, and Pakes (2004) and Bayer, Ferreira, and McMillan (2007) also combine micro level data with aggregate to estimate a demand models for differentiated products.

\textsuperscript{5}The idea is similar to the one proposed by Hoxby and Weingarth (2006) although the reassignment process used by Wake County, that is exploited in that paper, primarily depended on student characteristics such as income and race. There is no evidence that the school district, that we study in this paper, used these criteria.

\textsuperscript{6}In that sense, our approach to identification has some similarities with the spatial regression discontinuity design advocated by Black (1999).
countywide decline in public school enrollment. Total enrollment fell from approximately 44,000 to 28,000 within a 15-year period. Many parents and administrators clearly did not fully anticipate the speed and extent of this decline. This view is also consistent with the fact that the district tried, but failed to close schools before 2005. This failure was largely due to local politics and the difficulty to agree on a specific set of actions. This is not surprising since there was much disagreement about the causes of the decline in the local media. The actual school closing that took place in 2005/06 also came as somewhat of a surprise to many parents and administrators in the district. The changes were primarily due to the hiring of a new, energetic superintendent who convinced the school board to vote on his school closing plan in an up-or-down vote without allowing for amendments of his proposal.

The extent of the transformation of the school district implied significant changes in the administration, reassignments of students, teachers, and principals, as well the reconstitution of some schools. It is, therefore, likely that unobserved school characteristics also significantly changed during the period that we study.\textsuperscript{7} Our identification strategy implicitly relies on the fact that adjustments in residential housing markets are slower and more costly than adjustments of school choices. While it is not difficult for parents to move children from one public school to another within a district that practices a de facto open enrollment policy, it is more costly for parents to relocate in response to a shift in the distribution of school quality, especially if these shifts are not easily anticipated. Note that our model of school choice is consistent with this view and models parents as myopic. It only requires parents to correctly forecast the composition of the school in the next period. Under these plausible assumptions, peer effects predicted by the school assignment rules are not likely to be correlated with the actual unobserved

\textsuperscript{7}This insight also suggests that lagged endogenous variables may be useful instruments and we explore these ideas as part of our sensitivity analysis.
characteristics of most schools. As we will discuss below, our identification also seems to work in practice, and produces reasonable estimates of the magnitude of peer effects that are broadly consistent with other estimates reported in the literature.

Our empirical analysis is based on administrative panel data from one urban district, which prefers to remain anonymous. One key advantage of our data set is that the district provides transportation for all students in the district. As a consequence, we also have data from students living in the district that do not attend public schools. This allows us to model the extensive margin of the school choice problem. We use a short panel that includes observations from four school years, 2004-05 through 2007-08. Our analysis exploits the fact that our district implemented a plan to close a substantial number of underperforming schools across the district during the time period. As a result of the closures, the number of elementary and middle schools declined by approximately 25 percent following the 2005/06 school year. Importantly, we can reconstruct the school quality measures that were used by the district to determine which school to close. We can, therefore, closely approximate the objective function of the school district used in practice to close schools following the 2005/06 school year.

Closing underperforming schools based on exogenous quality criteria, that do not account for possible resorting, can create significant displacement of students. This in turn can induce exit of students from a district, aggravating problems of declining enrollment. This raises the question whether the district can avoid losing or displacing large number of students while reducing excess capacity. We, therefore, explore alternative objectives such as maximizing retention or minimizing disruptions (i.e: minimizing the number of students in schools that are closed). We find that these objectives imply closing a different set of schools. Some of the difference arises because our estimated unobserved school characteristics are only weakly correlated with the quality measures used by the district. This underlines the importance of incorporating consideration of parental perceptions of
school quality, and associated student sorting, in school closing decisions.

We also consider the objective of reducing stratification of students by demographic characteristics across schools. Here our findings highlight tradeoffs in setting school district policy. Our results suggest that superintendents confront a difficult dilemma: pursuing an equity objective, such as limiting demographic stratification across schools, results in the exit of many more students than are lost by an objective such as maximizing retention of students in the district or minimizing the number of students in schools chosen for closure.

achievement across demographic groups. Similarly, Neilson (2013) estimates a model of sorting across public and private schools using data from Chile. The main differences between the last two papers and our paper is our treatment of the supply side. We focus on managing school district capacity, which gives rise to a new game theoretic formulation of equilibrium in the market for public education.

Our approach is also related a more recent literature that estimates equilibrium models of educational markets. Epple and Sieg (1999) consider sorting of households in a metropolitan market and show that households stratification by income is primarily driven by difference in school spending across districts. Ferreyra (2007) estimates a similar model to study the impact of vouchers on sorting and student achievement. Finally, our work is also related to research on controlled choice mechanisms such as Abdulkadiroglu and Sonmez (2003), Abdulkadiroglu, Pathak, and Roth (2005, 2009), and Kesten (2010). These paper approach the school choice problem as a mechanism design problem. The game that we formulate can be interpreted as giving rise to individual rationality and self-selection constraints. Given that our focus is on managing district capacity, we explore a different set of research questions.

The rest of the paper is organized as follows. Section 2 discusses the data used in the empirical analysis. Section 3 develops our sequential game of managing school quality and sorting among possible school configurations. Section 4 discusses the estimator for the parameters of the model. Section 5 provides the parameter estimates. Section 6 discusses the main findings of our analyses that are based on the quantitative simulations of various school closing scenarios. Section 7 offers conclusions and discusses future research.
2 Data

Similar to many urban school districts, the district that we study in this paper has been experiencing declining enrollment and associated fiscal pressures have necessitated closing of schools. Since the district prefers to remain anonymous, we will refer to our district for the rest of the paper as the Center City School District (CCDS). The county in which CCSD is located contains more than 40 suburban school districts and is home to approximately sixty percent of the population of the metropolitan area. The county thus serves as a natural point of reference for summarizing the fortunes of CCSD relative to suburban school districts.

Figure 1 plots the student enrollment “market share” of the urban district considered in this paper relative to the broader educational market measured by all districts in the county. The district was maintaining its student share during the 1990’s when enrollment was rising in the market, but its enrollment and market share both dropped rapidly when metropolitan enrollment began to decline. Countywide births started to decline in the early 1990’s. Much of this decline in births was a result of the end of the “Echo Boom.” Largely as a result of the decline in births beginning in the early 1990’s, countywide enrollment began to decline in 1998. The district not only shared in the countywide decline in the student population, but experienced a disproportionate decline as its share of county enrollment also fell. To explore this phenomenon, we ranked school districts in the county by income and aggregated them into quartiles of roughly equal enrollment. We found that more affluent districts more than held their ground during the overall decline in enrollments. This finding is consistent with the notion that more affluent households exited the city and moved up the school district income hierarchy (“voting with their feet”). As a consequence, the district bore 75 percent of the countywide decline in public school enrollment. These declines created large excess capacities in the district.
Figure 1: Market Share of the Urban District

CCSD Enrollment and "Market Share"

CCSD Enrollment as Percent of County Enrollment (Right Scale)

Year

CCSD Enrollment (Left Scale)
In 2005, CCSD launched a series of initiatives aimed at addressing these challenges. This led to closure at the end of the 2005-06 school year of 22 schools serving students in Kindergarten through eighth grade.

The empirical analysis of this paper focuses on the impact of school closings on parental choices in CCSD. Our sample consists of all K-8 students that are part of the CCSD database between 2004 and 2007 and attended public schools for at least one year in 2004-2007. We thus exclude from our sample all private school students that never attended a public school. This is done for two reasons. First, our analysis of transitions between private and public schools shows that few students that attend private schools return to public schools. Second, we do not observe test scores for students that never attended public schools and, thus, must impute achievement for these students. We also eliminate high school students from our sample since high schools were not affected by the school closing plan adopted in 2005.

Table 1 provides summary statistics of student characteristics for our sample. The variables used in this study are defined as follows. For the moving indicator, we count as “moved” any student who attended a different school than in the previous year. This categorization includes students moving into the district, and students changing from elementary to middle school. We do this, as students moving from elementary to middle school face similar sorts of moving costs as those switching schools in another grade, such as acclimation to new facilities, teachers, and peers. The driving times variable denotes the median driving time from home to school in minutes. Individual achievement is measured as an average of all observed standardized test Z-scores.

In order to measure school capacity, we use the combination of two sources. First, the district gave us a time-invariant measure of capacity, which we call “stated capacity”. However, for some schools in some years, this stated capacity measure is lower than even actual enrollment. For this reason, we find the maximum actual enrollment for each school

11
Figure 2: Capacity the Urban District

Enrollment/Capacity By School: 2004

Enrollment/Capacity By School: 2005

Enrollment/Capacity By School: 2006

Enrollment/Capacity By School: 2007
Table 1: Summary Statistics of the CCSD Sample K-8

<table>
<thead>
<tr>
<th></th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>24660</td>
<td>24876</td>
<td>24489</td>
<td>23735</td>
</tr>
<tr>
<td>Sample Size: Public + Private</td>
<td>22158</td>
<td>21189</td>
<td>20333</td>
<td>19045</td>
</tr>
<tr>
<td>Sample Size: Public</td>
<td>21239</td>
<td>20180</td>
<td>19189</td>
<td>17956</td>
</tr>
<tr>
<td>Free or Reduced Lunch</td>
<td>17438</td>
<td>17840</td>
<td>17446</td>
<td>16866</td>
</tr>
<tr>
<td>Race: Black</td>
<td>14256</td>
<td>14219</td>
<td>13866</td>
<td>13312</td>
</tr>
<tr>
<td>Race: White</td>
<td>8498</td>
<td>8545</td>
<td>8405</td>
<td>8164</td>
</tr>
<tr>
<td>Race: Other</td>
<td>1906</td>
<td>2112</td>
<td>2218</td>
<td>2259</td>
</tr>
<tr>
<td>Moving choices indicator</td>
<td>10938</td>
<td>9920</td>
<td>12255</td>
<td>8796</td>
</tr>
<tr>
<td>Driving times to school attended</td>
<td>3.086</td>
<td>3.130</td>
<td>3.273</td>
<td>3.362</td>
</tr>
<tr>
<td>Individual achievement measure</td>
<td>-0.042</td>
<td>-0.033</td>
<td>-0.031</td>
<td>-0.045</td>
</tr>
</tbody>
</table>

over the years 2002-2007 and denote this magnitude as the “observed capacity.” Then, to create the capacity measure utilized in the optimal closing analysis, we simply take the maximum of stated and observed capacity.\(^8\) Using these capacity measures we find that the district had a capacity of 34,053 in 2005 before the school closing in all K-8 schools. The capacity was reduced to 24,588 after the school closings were implemented. For grades K-5 (6-8) the capacity was reduced from 13,192 (20,861) to 9,037 (15,550). Figure 2 plots the empirical distribution of capacity utilization – the fraction of enrollment over capacity – for the schools in the district. We find that only a few schools operated near capacity before the reforms. Even after the reforms, there was still some excess-capacity left in the system.

\(^8\)It will sometimes be necessary to calculate the capacity available for middle school students (grades 6-8). For schools serving kindergarten through eighth grade (K-8 schools), the 6-8 capacity measure is calculated as the overall capacity multiplied by the proportion of students in grades 6-8 in that year.
Finally, we have access to the quality measures that were used by the district to close schools. The district used a School Performance Index (SPI) in order to determine which schools to close. The goal of the SPI is to measure “the school’s contribution to student achievement”, and is based on a combination of regression specifications involving student test scores as the dependent variable.\textsuperscript{9} The SPI is a categorical ranking from the set \{1, 2, 3, 4\}, with 1 being the worst and 4 being the best. Using this categorical measure, CCSD closed 22 schools (primarily those with $SPI = 1$). To implement our counterfactual school closing analysis, we convert these categorical variables into continuous variables by regressing them on the average achievement measures reported. We consider the predicted SPI as the school quality ranking implied by the CCSD’s objective function when deciding which schools to close.

\section*{3 Managing School District Capacity}

We consider a sequential game in extensive form between a superintendent of a school district and a continuum of parents. In the first stage of the game, the superintendent determines the set of schools that are operational. In the second stage, parents enroll their children in one of the public schools or choose one of the outside options, taking into consideration that peer characteristics of schools are endogenous. Parents have idiosyncratic shocks to preferences for schools. These shocks are private information. The equilibrium concept that we use is Perfect Bayesian Equilibrium (PBE). We start by analyzing the second stage of the game and then consider the first stage using a standard backward induction argument.

\textsuperscript{9}Neilson (2013) uses a similar approach to measure school quality.
3.1 The Second Stage: Feasible Configurations & Peer Effects

Let $J_t$ denote the set of potential public schools that are available at time $t$.\textsuperscript{10} A school configuration is given by a set of schools that are active. Let $J_t^o \subseteq J_t$ denote a potential school configuration. Similarly, let $J_t^c$ denote the set of schools that are to be closed. By construction, we have $J_t = J_t^o \cup J_t^c$.

We make two assumptions regarding admission decisions of the district. First, we assume that the district operates an open enrollment system. Any student in the district can choose to enroll in any one of the available schools. Second we abstract from capacity constraints in the baseline model. For any configuration of schools, there are no binding capacity constraints at the school level. We relax this assumption below, introduce school level capacity constraints and discuss the impact of rationing.\textsuperscript{11}

The set of outside options is denoted by $O_t$. The choice set is thus given by $J_t^o \cup O_t$.\textsuperscript{12} Parents exercise school choice by enrolling their children in one of the public schools or opt out of the public school system.

\textsuperscript{10}Technically speaking the number of school options available to a student depend on the student’s grade. We account for this in the empirical analysis, but suppress this dependence for notational convenience here.

\textsuperscript{11}In our application, almost all schools in the district that were open had excess capacity. See Figure 2. During the sample time period, there were a small number of selective magnet programs that were operating at capacity and used lotteries to determine admissions. In our empirical analysis we found that the findings do not depend on our treatment of these schools. The results are similar if we include or exclude students attending oversubscribed magnet programs.

\textsuperscript{12}In our application, we assume that each student can attend a (generic) charter school, three different parochial school types, or an independent private school within the district. We do not model heterogeneity in schools within these types. In addition, the student can leave the district to attend a (typically suburban) school outside the district.
$x_{jt}$, and exogenous characteristics, $\xi_{jt}$. Endogenous characteristics are those that depend on the outcome of the choice or sorting process. In our application, $x_{jt}$ includes measures of peer characteristics such as average achievement of students in the school, and measures of demographic variables such as the proportion of students eligible for subsidized lunch, proportions in different racial/ethnic groups. Exogenous characteristics are, for example, the quality of the principal and the teachers.\footnote{These are typically difficult to observe for the econometrician.}

When parents make decisions about enrolling their children into public schools, they hold beliefs about the school peer effects. In equilibrium, these beliefs have to be consistent with the optimal strategies of parents and the superintendent.

To endogenize school characteristics via student sorting, consider a student that lives in the district. Let $z_{it}$ be the observed vector of characteristics of student $i$ at time $t$. Variables in $z_{it}$ include the student’s achievement, race, free or reduced lunch status, as well as a measure of behavioral problems. Let $d_{ijt-1}$ be an indicator variable which is equal to one if student $i$ attended school $j$ in period $t-1$. Previous school choices matter in our model since transferring to a new school is costly.\footnote{Switching to a new school may be costly since it requires the student to adapt to new school rules and acclimate to new peers, facilities, and teachers. The key assumption is that only the last year matters and captures of the full history of past choices (a first-order Markov assumption). This allows us to estimate the model based on conditional choice probabilities as discussed in detail below.}

We assume that parents have private information that can be characterized by a vector of idiosyncratic choice specific shocks, denoted by $\epsilon_{jt}$, and a vector of preferences over endogenous characteristics, denoted by $\beta_i$. A parent-student pair is completely characterized by a vector of characteristics $(z_{it}, d_{it-1}, \beta_i, \epsilon_{it})$.\footnote{We consider the parent-student pair as a single decision-making agent, and will thus use ”parent” and ”student” interchangeably.} We assume that the utility function of student $i$ at time $t$ is additively separable in the idiosyncratic preference shocks
and can thus be written:

\[ U_i(x_t, \xi_t, z_{it}, d_{it-1}, \beta_i, \epsilon_{it}) = \sum_{j \in J^o \cup O_t} d_{ijt} [u(x_{jt}, \xi_{jt}, z_{it}, d_{it-1}, \beta_i) + \epsilon_{ijt}] \]  \hspace{1cm} (1)

Given beliefs about endogenous school characteristics, the optimal strategy is then to choose the school that maximizes utility given the set of available school options. From the perspective of the superintendent and other parents, each parent’s decision is random due to the existence of private information, despite the fact that each parent uses a pure strategy. Integrating out the private information yields conditional choice probabilities for each school option denoted by:

\[ Pr\{d_{jt} = 1|x^o_t, \xi^o_t, J^o_t, z_{it}, d_{t-1}\} \]  \hspace{1cm} (2)

Aggregating choices among all parents yields the aggregate market shares for each public school, denoted by \(s_j\), and the market share for each outside option.

A set of strategies for each parent implies an allocation \((J^o_t, x^o_t, \xi^o_t)\). This is an equilibrium of the second stage of the game, if and only the vector \(x^o_t\) satisfies the following consistency requirement:

\[ x^o_{jk} = \frac{\int z_k Pr\{d_{jt} = 1|x^o_t, \xi^o_t, z, d_{t-1}\} f(z, d_{t-1}) dz \, dd_{t-1}}{\int Pr\{d_{jt} = 1|x^o_t, \xi^o_t, z, d_{t-1}\} f(z, d_{t-1}) dz \, dd_{t-1}} \quad \forall k, j \]  \hspace{1cm} (3)

The equilibrium of the second stage of the game thus implies that the system of equations given by (2) and (3), that are internally consistent. Beliefs are confirmed in equilibrium.

Thus far we have ignored school-level capacity constraints. To obtain a more realistic model, we now assume that each public school has a capacity constraint that is equal to \(n_j\). Equilibrium in the second stage of the game requires that the school capacity constraints are satisfied:

\[ \int Pr\{d_{jt} = 1|x^o_t, J^o_t, z, d_{t-1}\} f(z, d_{t-1}) dz \, dd_{t-1} \leq n_j \quad \text{for all} \; j \in J^o; \]  \hspace{1cm} (4)
We need to incorporate these capacity constraints in the characterization and computation of equilibrium of the second stage of the game. There are a number of different ways of implementing rationing rules. We consider a rule that is based on the idea that the parents are charged a shadow price for admission given by $p_j$ if there is excess demand for a school. Parents who want their children to attend a school facing excess demand have to invest effort in making an early application to the school, in cultivating the principal, and in pursuing other activities to enhance the likelihood that their children will get admitted. These activities are costly. In equilibrium, the shadow price is set equal to the value which guarantees that the capacity constraint holds with equality. The reduction in demand occurs because the shadow price is deducted from the utility that parents obtain from having their child attend the school.\footnote{Alternatively, we could use a random rationing rule via lotteries. That case is discussed in Geyer and Sieg (2013) who show how to modify the demand system in the presence of lotteries and compute the endogenous probabilities that determine the outcome of the rationing process.}

To formalize these ideas, let us assume that public school $j$ has shadow admission price denoted by $p_j$. We measure the shadow price in the same units as the unobserved school characteristic. It enters additively into the utility function. We obtain:

$$U_i(x_t, p_t, \xi_t, z_{it}, d_{it-1}, \beta_i, \epsilon_{it}) = \sum_{j \in J_o \cup O_t} d_{ijt} \left[ u(x_{jt}, \xi_{jt} - p_{jt}, z_{it}, d_{it-1}, \beta_i) + \epsilon_{ijt} \right]$$

(5)

We say that a school configuration with rationing, $(J_o, x_o, \xi_o, p_o)$, is feasible if it is an equilibrium of the second stage of the game, i.e. if and only if it satisfies the following three conditions:

a) The vector of school peer characteristics $x^o_t$ is consistent with the implied student sorting:

$$x^o_{jk} = z_{jk} = \frac{\int z_k Pr\{d_{jt} = 1|x^o_t, \xi^o - p^o_t, z, d_{t-1}\} f(z, d_{t-1}) dz dd_{t-1}}{\int Pr\{d_{jt} = 1|x^o_t, \xi^o - p^o_t, z, d_{t-1}\} f(z, d_{t-1}) dz dd_{t-1}} \quad \forall k, j$$

(6)
b) The market share of each school $s_j$ is less than or equal to the capacity constraint, i.e $s_j \leq n_j$.

c) The shadow prices satisfy the complimentary slackness condition:

$$p^o_{jt}(n_j - s_j) = 0$$

with $p^o_{jt} \geq 0$. It is straightforward to show that this rationing device is efficient. It allocates the available capacity to those who value it most.

### 3.2 The First Stage: Optimal Configurations

Next consider the first stage of the game. To solve the superintendent’s problem, we need to make an assumption about the objectives of the district. First, we characterize the objective function of the district used in actuality. We then discuss alternative objectives below.

Managing school capacity is often done with the explicit goal of closing "underperforming" schools. As we discussed in Section 2, the district estimated school-specific, achievement-based quality measures, $q_j$, that were taken as attributes of the schools (i.e: did not depend on sorting by households). A good approximation of the school district’s decision problem is to assume that the district maximizes a weighted average of school quality:

$$Q = \sum_{j \in J^o} w_j q_j$$

where $w_j$ denotes the weight of school $j$. The district maximizes quality over the set of feasible school configurations that result in an equilibrium of the second stage of the game. Moreover, it faces the constraint that total excess capacity in the district does not exceed a given threshold:

$$\sum_{j \in J^o} (n_j - s_j) \leq c.$$
where $n_j$ is the capacity of school $j$ and $s_j$ the enrollment.

We define a Perfect Bayesian Equilibrium for the game. A set of beliefs and strategies, which imply an allocation, $(J^o_t, x^o_t, \xi^o_t)$, is a PBE of our game in extensive form, if and only if,
a) for each possible school configuration $J^o_t$ there exists a Bayesian Nash equilibrium in the second stage;
b) strategies that imply $(J^o_t, x^o_t, \xi^o_t)$ are an equilibrium for the second stage and satisfy the excess capacity condition in equation (9); and
c) Given all second stage equilibria given in (a), there does not exist another equilibrium that satisfies equation (9) and yields higher welfare for the superintendent.

The objective in (8) provides no role for considerations other than school performance that are potentially important to a district. First, parental school choices in response to school closures may exacerbate heterogeneity in peer quality among schools and/or demographic stratification within the district. Second, maximizing (8) ignores the dependence of choices on proximity to school. This may make it difficult to retain students, especially if some students are forced to commute longer distances. Maximizing (8) may also imply large reallocations of students among schools and, therefore, create significant transitional problems. Of course, in pursuing the objective in (8), the district was no doubt cognizant of these potential ramifications of school closings and likely made subjective adjustments to accommodate such considerations in making decisions about which schools to close. Our goal is to provide a framework to enable a superintendent to incorporate such considerations in a more systematic way.

To gain some more insights into these important issues, we consider three alternative objectives: limiting demographic stratification across schools, retaining students in the district, and minimizing the number of students attending schools chosen for closure. Balancing these different objectives captures the dilemma that the superintendent faces.
As we will see in the numerical analysis below, these objectives lead to quite different outcomes with regards to school closings.

To measure inequality in the provision of education, we use the sum of the weighted squared deviation between school j’s and the district’s characteristics over public schools. Let $\overline{z}^d$ denote the mean characteristics at the district level. Summing over all open public schools gives the total inequality index:

$$I = \sum_{j \in J^o} \sum_{k=1}^{K} \omega_k (\overline{z}_{jk} - \overline{z}^d_k)^2$$

where $\omega_k$ is the weight assigned to school characteristic $k$. \(^{17}\)

The district wishes to attract and retain students in district schools. A measure of student retention is given by:

$$R = \sum_{j \in J^o} \int Pr\{d_{jt} = 1|x_t, \xi_t, J_t^o, z, d_{t-1}\} f(z, d_{t-1}) \, dz \, dd_{t-1}$$

Note that the probability here denotes to the probability in the baseline model.

Finally, the district wishes to limit the number of students that are attending schools that are closed. A measure of those students is given by:

$$D = \sum_{j \in J^c} \int Pr\{d_{jt} = 1|x_t, \xi_t, z, d_{t-1}\} f(z, d_{t-1}) \, dz \, dd_{t-1}$$

where $p_{jt}^b$ denote the baseline probabilities associated with the full choice set. By studying the equilibrium properties of the game for alternative objective functions, we can characterize the trade-offs faced by the superintendent of a school district.

\(^{17}\)We determine the numerical values for the weights ($\omega_k$’s) in our computational analysis in order to scale the characteristics such that they have approximately equal weight in the objective function.
3.3 Properties of Equilibria

In this section we discuss existence and uniqueness of equilibrium of our model and discuss its properties. One useful result is that we can recast the problem of existence of equilibrium as a problem of finding an optimal solution to a planer’s problem. The next proposition formalizes this result.

Proposition 1 The problem of computing a PBE equilibrium for our game can be recast as a non-linear integer programming problem. In particular, there exists an equilibrium to the game if there exists an optimal solution to a well-defined non-linear integer programing problem. Equilibrium is unique if the solution to that non-integer programming problem is unique.

Proof:

We consider an algorithm that consists of two loops. For a given set of schools that remain open, i.e. a given set $J^o$, the inner loop determines a set of beliefs over school peer effects that are consistent with optimal parental sorting. We call such an allocation a feasible school configuration. A feasible configuration, by definition, corresponds to an equilibrium in the second stage of the game. The outer loop of the algorithm then solves the maximization problem of the superintendent. We show that this problem is a non-linear integer programming problem. A solution of the outer loop is, therefore, a PBE of the sequential game.

First consider the case in which we ignore school level capacity constraints, the inner loop is a standard learning algorithm that computes a fixed point of a mapping that maps the beliefs over peer qualities into realized peer equalities. Consider a possible combination of schools that remain open, $J^o$. At the beginning of iteration $n$, we have a vector of school characteristics denoted by $\tilde{z}_{jn}^n$, which characterize the beliefs of the individuals.
at iteration $n$. Given these school characteristics $\bar{z}_j^n$ we can evaluate the conditional choice probabilities using equation (2). We can then update the school characteristics and obtain $\bar{z}_j^{n+1}$ for each school $j \in J^o$ using equations (3). We then forward-iterate this mapping until convergence of school characteristics and conditional choice probabilities. A fixed point of this mapping is then by construction an equilibrium of the second stage of the game.

If there are binding capacity constraints, we also have to iterate on the shadow prices associated with each school. The basic idea here is that we need to increase the shadow prices of schools with excess demand and set the shadow prices to zero for schools with excess capacity. Again we iterate until convergence. A fixed point of this modified mapping is then by construction an equilibrium of the second stage of the game with binding capacity constraints.

The outer loop solves the planer's problem that is associated with the superintendent’s optimization problem. Once we have characterized a feasible allocation for each school configuration, we check whether the aggregate excess capacity constraints are met and evaluate the objective function of the superintendent. Since the set of schools is finite, there are only a finite number of possible combinations of schools. As a consequence, the algorithm computes the solution to an integer programming problem. It is a non-linear integer programming problem since the control variables enter the conditional choice probabilities and the endogenous peer qualities in a non-linear way. The outer loop of the algorithm then searches over all feasible school configurations and computes a solution to this non-linear integer programming problem. A solution to that problem is then, by construction, a PBE of the game in extensive form.

Q.E.D.
If we ignore school-specific capacity constraints, the second stage of the game is almost identical to the neighborhood sorting game developed in Bayer and Timmins (2005). The arguments presented in that paper can be easily extended to show that there exists an equilibrium in the second stage of the game. As discussed in detail in Bayer and Timmins (2005), existence of equilibrium without rationing follows from a straightforward contraction mapping argument. With capacity constraints, we also need to iterate on the shadow prices for each school that faces excess demand in equilibrium. While this modified mapping is not necessarily a contraction mapping, we find that our algorithm converges in practice for many reasonable specifications. We thus conclude that second stage equilibria exist and can be computed for the type of specification discussed in detail below.

While existence is not an issue, uniqueness is more complicated. As in most models with endogenous peer effects, there is some scope for multiplicity of equilibria. We can, however, rank second-stage equilibria for a given school configuration using the same objective function of the superintendent that we use to evaluate different school configurations. This suggests using the welfare ranking of equilibria as a refinement concept. Generically, there exists a unique maximum if the number of equilibria is finite. We, therefore, conclude that for each school configuration there often exists a unique PBE that maximizes the welfare of superintendent.

Since there are only a finite number of feasible school configurations, there exists, at least, one PBE for the game in extensive form. Uniqueness of equilibrium of the game in extensive form, follows from the same argument that implies uniqueness of equilibrium for the second stage equilibrium, i.e: given uniqueness in the second stage; it is trivial to show that the first stage district’s problem will be unique.

For reasonable problems, an exact solution can be determined by simply evaluating all possible combinations of school closings. The problem increases exponentially in com-
plexity with the number of schools to be closed, making enumeration of all alternatives an infeasible approach for large problems.\footnote{Future applications of our framework may be able to draw on ongoing work in operations research. One branch of work seeks to enhance speed and accuracy of methods for computing an approximate solution that is reasonably close to the optimum (Belotti, Lee, Liberti, Margot, and Wächter, 2009) while also providing a measure of the extent by which the approximate solution falls short of the full optimum. Our computational problem is in the domain of mixed-integer polynomial optimization, and within that domain, it is distinguished by having low-degree polynomials. A related branch of work focuses particularly on this class of problems (Burer and Letchford, 2012).}

\section{Estimation}

In order to numerically compute the equilibria discussed in Section 3, we need to parametrize the model and estimate its parameters. Parameters corresponding to parental preferences can be estimated using standard techniques developed in the demand literature of differentiated products.

Utility of individual $i$ in school $j$ in year $t$ is given by:

$$U_{ijt} = \sum_{k=1}^{K} x_{jkt} \beta_{ikt} + \xi_{jt} + \epsilon_{ijt}$$

(13)

where the $k^{th}$ characteristic of school $j$ is denoted by $x_{jkt}$ and

$$\beta_{ikt} = \alpha_{0k} + \sum_{m=1}^{M} \alpha_{1kl} z_{imt} + \sigma_k u_{ik}$$

(14)

and the $z_{imt}$ is the $m^{th}$ component of individual $i$’s characteristics at time $t$. The random coefficient errors are time invariant and satisfy: $u_{ik} \sim N(0,1)$.\footnote{We also estimated a model that allowed for non-zero correlations among the random coefficients and found no improvements. Alternatively, one could generate realistic substitution patterns in demand using a nested Logit model as shown by McFadden (1981) and Goldberg (1995).} Define the fixed effect
of school \( j \) in year \( t \) as:

\[
\delta_{jt} = \sum_{k=1}^{K} \alpha_{0k} x_{jkt} + \xi_{jt}
\]  

(15)

We can then write the school specific utility of individual \( i \) in year \( t \) as:

\[
u_{ijt} = \delta_{jt} + \sum_{k=1}^{K} \sum_{m=1}^{M} \alpha_{1km} x_{jkt} z_{ilt} + \sum_{k=1}^{K} \sigma_{k} x_{jkt} u_{ik} + \epsilon_{ijt}
\]  

(16)

For expositional simplicity, we have ignored travel times and shadow prices when writing the school-specific utility above. We have also omitted the moving costs which are given by \( mc_{it} = \gamma_{it} 1\{d_{jt} \neq d_{kt-1}\} \) where \( \gamma_{it} = \gamma_0 + \sum_{m=1}^{M} \gamma_{1l} z_{ilt} + \sigma_{k+1} u_{ik+1} \). We add these terms to the model specification when we estimate the model.

Idiosyncratic shocks in the utility function, \( \epsilon_{ijt} \), follow a Type I extreme value distribution (McFadden, 1974). Conditional on the observed characteristics \( (x_t, z_{it}, d_{it-1}) \) and the unobserved shock \( u_i \), the choice probabilities are then given by:

\[
Q_{ijt} = \exp\left(\delta_{jt} + \sum_{k=1}^{K} \sum_{l=1}^{L} \alpha_{1kl} x_{jkt} z_{ilt} + \sum_{k=1}^{K} \sigma_{k} x_{jkt} u_{ik}\right) / \sum_m \exp\left(\delta_{mt} + \sum_{k=1}^{K} \sum_{l=1}^{L} \alpha_{1kl} x_{mkt} z_{ilt} + \sum_{k=1}^{K} \sigma_{k} x_{mkt} u_{ik}\right)
\]  

(17)

In our application, we model choices for four consecutive time periods (\( T=4 \)):

\[
Pr\{d_{ij1}, d_{ik2}, d_{il3}, d_{im4}|x_t, z_{it}, u_i, d_{in0}\} = Q_{ij1} Q_{ik2} Q_{il3} Q_{im4}
\]  

(18)

Given that we do not observe \( u_i \), we need to integrate out the errors to obtain conditional choice probabilities that only depend on observables.

\[
Pr\{d_{ij1}, d_{ik2}, d_{il3}, d_{im4}|x_t, z_{it}, d_{in0}\} = \int Q_{ij1} Q_{ik2} Q_{il3} Q_{im4} dF(u_i)
\]  

(19)

Recall that the random coefficients are time invariant and normally distributed. We approximate the joint distribution of the random coefficients using quadrature methods (Skrainka and Judd 2011).

Estimation of the model proceeds in two stages. First, we estimate the discrete choice model with school-year level fixed effects. Second, we decompose these fixed effects into
observed and unobserved components using an IV strategy. Note that the conditional choice probabilities in equation (14) depend on the parameters $\alpha_1$, $\sigma$, and the mean utilities $\delta = (\delta_{11}, \ldots, \delta_{JT})$. We can, therefore, estimate these parameters using a ML estimator. The likelihood function is given by

$$L = \prod_{i=1}^{N} \int Q_{ij1} Q_{ik2} Q_{il3} Q_{im4} dF(u_i)$$

(20)

Before we proceed, we offer three observations. First, our likelihood function conditions on the observed average peer characteristics for each school. As we discussed in the modeling section, these peer characteristics are the result of student sorting based on unobservables, and are, therefore, potentially endogenous. Furthermore, there is scope for multiplicity of equilibria in these types of sorting games. By conditioning on the observed peer characteristics for schools, we are explicitly conditioning on the equilibrium under which the data were generated. Recall that we consider an economy with a continuum of individuals, each one of whom has a negligible impact on the peer characteristics of a school. The estimator only exploits that each individual behaves optimally given the observed collective choices made by other individuals, i.e. given the peer characteristics of each school. As a consequence, we can consistently estimate the parameters of the model. Second, since we are conditioning on observed equilibrium outcomes in the first stage, we can implement the estimator without having to compute the equilibrium of the model. Finally, we can identify and estimate the parameters of the likelihood function by appealing to large $N$ (number of students), finite $J$ (number of schools) asymptotics.

The school specific fixed effects are function of the observed and unobserved school characteristics.

$$\delta_{jt} = \alpha_0 x_{jt} + \xi_{jt}$$

(21)

Bayer and Timmins (2005) provide a comprehensive Monte Carlo study which shows that these types of estimators are well behaved in large samples. See also Bayer, McMillan, and Reuben (2004).
We assume this error term is uncorrelated over schools and time.\textsuperscript{21} Following Berry (1994), we assume that $E[\xi_{jt} \mid w_{jt}] = 0$ for some instruments $w_{jt}$.\textsuperscript{22} The main difficulty encountered in this stage of the estimation is that most observed school characteristics are peer measures and, therefore, endogenous outcomes of the sorting process. We, therefore, use an instrumental variable approach to account for the endogeneity of peer characteristics. We use administrative school assignment rules, that are implemented by the school district, to predict the composition of each school before and after the school closing. The validity of these instruments relies on the assumption that peer measures predicted by administrative assignment rules are not correlated with unobserved school characteristics. One can justify this assumption by noting that the school assignment rules used by our district are primarily a function of distance, i.e. the district tends to assign a student to the school that is closest to the home of residence. The predicted compositions of schools can, therefore, be viewed as nonlinear transformation of historical, spatial residential sorting patterns. In that sense, our approach to identification has some similarities with the spatial regression discontinuity design advocated by Black (1999).\textsuperscript{23}

We have shown in Section 2 of this paper that our school district experienced large changes prior and after the school closings. These changes were largely caused by a raptly shrinking district wide enrollment. During the period leading up to the school closing, the district bore 75 percent of the countywide decline in public school enrollment. We are thus studying the opposite of case that was analyzed in Hoxby and Weingarth (2006). One can then ask the question what economic and behavioral models of school and resi-

\textsuperscript{21}This assumption is primarily so that we can use lagged peer effects as instruments. For the Hoxby-Weingarth instruments, we can correct for heteroskedastic and autocorrelated errors.

\textsuperscript{22}Berry, Linton and Pakes (2004) discuss the asymptotic properties of the second stage estimator. The basic requirement is that $N$ grows fast enough so that the ratio between $J \ln(J)/N$ goes to zero. In our application that ratio is approximately 0.069.

\textsuperscript{23}Here the idea is that school district boundaries create spatial variation in observed school quality that is uncorrelated with unobserved local neighborhood characteristics.
dential choice are broadly consistent with our proposed instrumental variable approach. First, it seems necessary to assume that many parents and administrators clearly did not fully anticipate the speed and extent of this decline and its impact on unobserved school characteristics. We argued above that this is not an implausible assumption. Second, we implicitly assume that unobserved school characteristics underwent significant changes during the period that we study. These changes were caused by turn-over in the administration, reassignments of teachers and principals, as well the reconstitutions of some of the schools. None of these are easily measured or observed by the econometrician.\footnote{This insight also suggests that lagged endogenous variables may be useful instruments and we explore these ideas as part of our sensitivity analysis. These instruments are in the spirit of work in the estimation of dynamic panel data models which uses timing assumptions and lagged endogenous variables to generate instrumental variables (Arellano and Bond, 1991). We implement this estimator below as a robustness check.} Third, our identification strategy relies on the fact that adjustments in residential housing markets are slower and more costly than adjustments of school choices. While it is not difficult for parents to move children from one public school to another within a district that practices a de facto open enrollment policy, it is much more costly for parents to relocate in response to a shift in the distribution of school quality, especially if these shifts are not easily anticipated. There is much evidence in the literature that supports this assumption.\footnote{Epple, Romano and Sieg (2012) discuss the magnitude of relocation within metropolitan areas and the magnitude of the costs associated with such relocation.} Finally, parents cannot be too sophisticated in their forward looking school choice behavior. Our model of school choice is consistent with view and models parents as somewhat myopic. It only requires parents to correctly forecast the composition of the school in the next period. Under these four assumptions, peer effects predicted by the school assignment rules are not likely to be correlated with the actual unobserved characteristics of most schools.
5 Estimation Results

We estimate our demand models with and without school-year fixed effects. We also implement the fixed effect Logit model with and without random coefficients. For computational reasons, we only implement the random coefficients model for a 10 percent random subsample of the original sample. For comparison purposes we also estimate the fixed effects Logit model without random coefficients using the 10 percent sample. There are 31,684 unique students in the full sample and 3,304 in the random subsample. Tables 3-4 in Appendix A of the paper report the first stage parameter estimates and estimated standard errors. We find that the fixed effects model fits the data much better than the model without fixed effects. The point estimates are similar using the full and the 10 percent sample. Adding random coefficients to our specification improves the fit of the model. Table 2 summarizes the point estimates and estimated standard errors for different specifications of our demand model. Note that all of these coefficients are identified based on the observed conditional choice probabilities and therefore do not rely on the validity of our instrumental variables.

We then implement the second stage using the different IV techniques as discussed above. Table 4 reports the point estimates and estimated standard errors for the second stage estimator. We consider three different model specification. First, we report OLS estimates which do not account for the endogeneity of the peer characteristics. Second, we use a simple panel data estimator that uses lagged endogenous peer characteristics. We primarily report these estimates as robustness checks. Our preferred model specification is based on the predicted peer levels based on the assignment rules. We denote these instruments as the Hoxby & Weingarth instruments.
Table 2: First Stage Estimates

<table>
<thead>
<tr>
<th></th>
<th>No Fixed Effects</th>
<th>Fixed Effects</th>
<th>Fixed Effects</th>
<th>Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Random Coef</td>
<td>No Random Coef</td>
<td>No Random Coef</td>
<td>Random Coef</td>
</tr>
<tr>
<td></td>
<td>Full Sample</td>
<td>Full Sample</td>
<td>10 % Sample</td>
<td>10 % Sample</td>
</tr>
<tr>
<td>School Ability × ability</td>
<td>1.3433 0.0374</td>
<td>1.3715 0.039</td>
<td>1.4848 0.1240</td>
<td>1.8723 0.1609</td>
</tr>
<tr>
<td>School Ability × black</td>
<td>-0.1861 0.0647</td>
<td>-0.2908 0.073</td>
<td>-0.1671 0.2475</td>
<td>-0.6472 0.2781</td>
</tr>
<tr>
<td>School Ability × FRL</td>
<td>0.5255 0.0651</td>
<td>0.6083 0.072</td>
<td>0.5990 0.2445</td>
<td>0.3268 0.2768</td>
</tr>
<tr>
<td>School Black × ability</td>
<td>0.7657 0.0368</td>
<td>0.6855 0.0428</td>
<td>0.7710 0.1423</td>
<td>0.8278 0.1598</td>
</tr>
<tr>
<td>School Black × black</td>
<td>0.7657 0.0368</td>
<td>3.0924 0.0850</td>
<td>3.1558 0.2876</td>
<td>3.6778 0.3044</td>
</tr>
<tr>
<td>School Black × FRL</td>
<td>-0.6481 0.0660</td>
<td>-0.6243 0.0825</td>
<td>-0.8110 0.2794</td>
<td>-0.7934 0.2906</td>
</tr>
<tr>
<td>School FRL × ability</td>
<td>0.4941 0.0391</td>
<td>0.3075 0.0486</td>
<td>0.4559 0.1639</td>
<td>0.4105 0.1838</td>
</tr>
<tr>
<td>School FRL × black</td>
<td>-1.8534 0.0716</td>
<td>-0.8734 0.0839</td>
<td>-1.3920 0.2844</td>
<td>-1.0795 0.3212</td>
</tr>
<tr>
<td>School FRL × FRL</td>
<td>2.4717 0.0777</td>
<td>3.9485 0.0860</td>
<td>4.2640 0.2875</td>
<td>5.031 0.3502</td>
</tr>
<tr>
<td>School FRL × suspensions</td>
<td>0.1143 0.0064</td>
<td>0.1359 0.0095</td>
<td>0.1409 0.0282</td>
<td>0.1659 0.0373</td>
</tr>
<tr>
<td>School suspensions × ability</td>
<td>0.0150 0.0065</td>
<td>0.0293 0.0101</td>
<td>0.0423 0.0298</td>
<td>0.0244 0.0406</td>
</tr>
<tr>
<td>School suspensions × black</td>
<td>-0.1282 0.0136</td>
<td>-0.0615 0.0203</td>
<td>-0.0620 0.0622</td>
<td>-0.1145 0.0788</td>
</tr>
<tr>
<td>School suspensions × FRL</td>
<td>0.0809 0.0144</td>
<td>0.0613 0.0204</td>
<td>0.0830 0.0615</td>
<td>0.1179 0.0824</td>
</tr>
<tr>
<td>School suspensions × suspensions</td>
<td>0.0154 0.0010</td>
<td>0.0320 0.0012</td>
<td>0.0334 0.0035</td>
<td>0.0452 0.0053</td>
</tr>
</tbody>
</table>
Table 3: First Stage Estimates (cont)

<table>
<thead>
<tr>
<th></th>
<th>No Fixed Effects</th>
<th>Fixed Effects</th>
<th>Fixed Effects</th>
<th>Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No Random Coef</td>
<td>No Random Coef</td>
<td>No Random Coef</td>
<td>Random Coef</td>
</tr>
<tr>
<td></td>
<td>Full Sample</td>
<td>Full Sample</td>
<td>10 % Sample</td>
<td>10 % Sample</td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Moving costs</td>
<td>-4.4206</td>
<td>0.0296</td>
<td>-4.3259</td>
<td>0.0241</td>
</tr>
<tr>
<td>Moving costs × ability</td>
<td>-0.2953</td>
<td>0.0186</td>
<td>-0.2766</td>
<td>0.0136</td>
</tr>
<tr>
<td>Moving costs × black</td>
<td>-0.0736</td>
<td>0.0327</td>
<td>-0.1650</td>
<td>0.0250</td>
</tr>
<tr>
<td>Moving costs × FRL</td>
<td>0.1992</td>
<td>0.0333</td>
<td>0.1882</td>
<td>0.0263</td>
</tr>
<tr>
<td>Moving costs × suspensions</td>
<td>0.0421</td>
<td>0.0039</td>
<td>0.0500</td>
<td>0.0027</td>
</tr>
<tr>
<td>Travel costs</td>
<td>-0.3716</td>
<td>0.0041</td>
<td>-0.4223</td>
<td>0.0033</td>
</tr>
<tr>
<td>Travel costs × ability</td>
<td>-0.0017</td>
<td>0.0029</td>
<td>0.0003</td>
<td>0.0017</td>
</tr>
<tr>
<td>Travel costs × black</td>
<td>0.1118</td>
<td>0.0046</td>
<td>0.1092</td>
<td>0.0034</td>
</tr>
<tr>
<td>Travel costs × FRL</td>
<td>-0.1226</td>
<td>0.0047</td>
<td>-0.1217</td>
<td>0.0035</td>
</tr>
<tr>
<td>School Ability × Standard Normal</td>
<td></td>
<td></td>
<td></td>
<td>2.411</td>
</tr>
<tr>
<td>School FRL × Standard Normal</td>
<td></td>
<td></td>
<td></td>
<td>2.4307</td>
</tr>
<tr>
<td>School Black × Standard Normal</td>
<td></td>
<td></td>
<td></td>
<td>1.645</td>
</tr>
<tr>
<td>School Susp × Standard Normal</td>
<td></td>
<td></td>
<td></td>
<td>0.4159</td>
</tr>
<tr>
<td>Likelihood</td>
<td>160,902</td>
<td></td>
<td>136,777</td>
<td>13,854</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13,702</td>
</tr>
</tbody>
</table>
Table 4: Second Stage Estimates

<table>
<thead>
<tr>
<th></th>
<th>No Random Coefficients</th>
<th>Random Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>Lagged Regressors</td>
</tr>
<tr>
<td></td>
<td>(.457)</td>
<td>(.507)</td>
</tr>
<tr>
<td>School Black</td>
<td>-1.713</td>
<td>-1.707</td>
</tr>
<tr>
<td></td>
<td>(.163)</td>
<td>(.168)</td>
</tr>
<tr>
<td>School Ability</td>
<td>-.441</td>
<td>.061</td>
</tr>
<tr>
<td></td>
<td>(.279)</td>
<td>(.292)</td>
</tr>
<tr>
<td>School Suspensions</td>
<td>-.171</td>
<td>-.145</td>
</tr>
<tr>
<td></td>
<td>(.031)</td>
<td>(.032)</td>
</tr>
</tbody>
</table>

All estimators control for elementary school and middle school fixed effects.

Overall we find that most parameters are estimated with good precision and that the sign of the parameters are plausible. We plot the implied distributions of the random coefficients in Figure 3-6 which are shown in the appendix. The vast majority of students in the sample value schools with high student achievement. High achieving students have a higher valuation of mean achievement than low achieving students. The valuation of the peer effects measured by school-level proportions of Black and FRL students differs significantly by individual characteristics. Both distributions are bimodal. Students on FRL tend to prefer schools with high FRL proportions and vice versa. Similar results hold for race. Our findings, therefore, support the similarity hypothesis; individuals prefer environments that allow them to interact with individuals that are similar in observed
characteristics.

The impact of school-level mean number of suspensions is ambiguous. High suspensions in a school may be an indicator of a school with troubled students, which is an undesirable school characteristic. On the other hand, it may indicate that a no-excuses approach to disruptive behavior is enforced by the administration, which could be positively valued by some parents. We also find that parents are reluctant to change schools once they have made their initial choices. This is reflected in our estimates of the moving costs, which are identified by the lagged observed choices made by students. Similarly, students also face significant travel costs. School choices that force students to commute for longer times are not popular.

Having estimated the parameters of our demand model, we can determine whether the quality measures used by the district are similar to the measures used by parents when making school choice decisions. A simple test of this hypothesis is to determine whether closed schools are perceived by parents to be less desirable than schools that remained open. We perform an unconditional test based on the difference of the first stage school-year fixed effects among the two type of schools. Implementing this test we find that the mean of the fixed effects for closed schools is significantly lower than the mean of the fixed effects of schools that remained open. However, the “observed quality” reflects differences in peer characteristics across schools. We, therefore, perform a second test that conditions on observed peer measures and allows for potential time trends. To implement this approach we adopt a difference-in-difference strategy in the second stage of the demand estimation procedure. We implement this test by adding the following variables to our second stage: a) a post reform time dummy; b) an indicator for a school that was closed.

Overall, we do not find much evidence that the downsizing of the district affected parents’ perceptions of the distribution of school qualities. Once we condition on observed
peer characteristics, closed schools are not systematically worse in the perception of parents than the schools that remained open (i.e: parents perceive both sets of schools as having similar unobserved school qualities). This reinforces the expectation that that student sorting will be a key consideration in determining which schools to close.

6 School Closing Analysis

In our application, the district closed underperforming schools using a ranking of schools by a unidimensional performance measure. One key advantage of this approach is that it is straightforward to solve the optimization of the school district even for a large number of schools to be closed. The optimal policy is a cut-off rule; schools are closed in inverse order of ranking until the required capacity reduction is achieved. We can compare outcomes implied by this closing rule with outcomes under alternative objectives to characterize the trade-offs faced by a superintendent.26

The ideal comparison would be to compare the school closing agenda that was implemented by the district with school closing scenarios predicted by our model under alternative objective functions. However, considering all possible school configurations rapidly becomes infeasible as the number of schools that can potentially be closed increases. Due to this curse of dimensionality, we cannot compute equilibria for our model on the scale implemented by the CCSD. Instead we solve a different problem that is computationally feasible and mimics the problem faced by the district, albeit on a smaller scale. We consider the case in which the district closes the three lowest ranked schools

26 As we noted previously, the superintendent very likely considered more subjective factors in making closing decisions, and those considerations may have resulted in departure from strict adherence to the performance ranking of schools. Hence, our analysis of the "actual" closing policy abstracts from the subject judgments that are an inherent part of a superintendent’s decision making responsibilities.
(using the district measure of school quality). We compare this "actual" scenario with closing down (at least) the same level of capacity under alternative objective functions.\textsuperscript{27}

Table 5 summarizes the key findings of this paper. The table refers to solutions of the optimal closing problem using the four different objective functions: 1) maximizing school quality using the quality measures of the district, 2) minimizing differences in peer characteristics across schools, 3) minimizing number of students leaving the district, i.e maximizing retention; 4) minimizing the number of students who relocate as a result of the school closings, i.e. minimizing dislocation. As a shorthand, we refer to these objectives as Quality, Similarity, Retention, and Dislocation objectives. We solve the model with and without imposing school-level capacity constraints.\textsuperscript{28}

We begin with the cases that ignore the school-level capacity constraints. For ease of reference, we will refer to elements of Table 5 by column number and row letter. All results are predictions of the model, so we drop "predicted" as a modifier in the following summary for expositional ease. The results in (1a) show, for the Quality objective, that the enrollment in schools selected for closure is 619 students. Schools chosen for closure under this objective have peer characteristics (1b to 1d) that differ somewhat from the district averages (9b to 9d), having lower achievement, more low-income students, and more black students. We also see from (1e) that these closings induce net departure from district schools of 389 students.

\textsuperscript{27}To reduce the computational burden, all computations are based on our estimated demand model with school-year fixed effects, but without random coefficients. Allowing for random coefficients in the demand model significantly increases the complexity of computing equilibria, but does not seem to affect the qualitative and quantitative implications of our model.

\textsuperscript{28}It is useful to translate the results in number of students affected by the policies. To accomplish this task, recall that there were 8,245 students in grades 6-8 in the district in 2005. We, therefore, use this number to translate changes in market shares into "number of students."
Table 5: Optimal School Closing Analysis

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Capacity Constraint</td>
<td>Capacity Constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a Enroll. Closed Schools</td>
<td>619</td>
<td>1065</td>
<td>764</td>
<td>619</td>
<td>619</td>
<td>760</td>
<td>723</td>
<td>619</td>
<td>8217</td>
</tr>
<tr>
<td>b Mean FRL</td>
<td>0.737</td>
<td>0.702</td>
<td>0.714</td>
<td>0.737</td>
<td>0.737</td>
<td>0.726</td>
<td>0.715</td>
<td>0.737</td>
<td>0.707</td>
</tr>
<tr>
<td>c Mean Black</td>
<td>0.653</td>
<td>0.615</td>
<td>0.583</td>
<td>0.653</td>
<td>0.653</td>
<td>0.648</td>
<td>0.606</td>
<td>0.653</td>
<td>0.599</td>
</tr>
<tr>
<td>d Mean Achievement</td>
<td>-0.054</td>
<td>-0.032</td>
<td>-0.013</td>
<td>-0.054</td>
<td>-0.054</td>
<td>-0.066</td>
<td>-0.006</td>
<td>-0.054</td>
<td>-0.038</td>
</tr>
<tr>
<td>e Students Leaving CCSD</td>
<td>389</td>
<td>583</td>
<td>296</td>
<td>389</td>
<td>390</td>
<td>468</td>
<td>310</td>
<td>390</td>
<td></td>
</tr>
<tr>
<td>f Mean FRL</td>
<td>0.740</td>
<td>0.711</td>
<td>0.725</td>
<td>0.740</td>
<td>0.740</td>
<td>0.729</td>
<td>0.728</td>
<td>0.740</td>
<td></td>
</tr>
<tr>
<td>g Mean Black</td>
<td>0.655</td>
<td>0.628</td>
<td>0.571</td>
<td>0.655</td>
<td>0.655</td>
<td>0.656</td>
<td>0.611</td>
<td>0.655</td>
<td></td>
</tr>
<tr>
<td>h Mean Achievement</td>
<td>-0.055</td>
<td>-0.050</td>
<td>-0.011</td>
<td>-0.055</td>
<td>-0.055</td>
<td>-0.0677</td>
<td>-0.001</td>
<td>-0.055</td>
<td></td>
</tr>
</tbody>
</table>
Relative to the district schools and the schools that are closed, these “leavers” have similar characteristics as the students in the closed schools. These results illustrate the challenge that the district faces in attempting to close underperforming schools. Pursuit of this objective displaces students that attend schools that are closed and induces flight of students from the district. As this column shows, the loss of students from the district amounts to approximately 60 percent of the number of students in the schools that were closed. Not all of these departures are by students who attended schools that were closed. Inflow of students from closed schools results in changes in peer characteristics of schools that remain open, and those changes in peer characteristics induce some incumbent students in those schools to exit the district.

Column (2) reports the results with the Similarity objective. Note that we can interpret this objective as minimizing peer quality differences among schools. The result shows that this approach is very costly. Not only does it affect a larger number of students that are affected by closings (1065), but it also induces a much larger flight of students from the district.

Column (3) reports results with the Retention objective. This objective results in closing of schools attended by a larger number of students (764) than the Quality objective. Relative to the district averages (9b to 9d), these schools have higher achievement (2d) and are attended by fewer low-income (2b) and black (2c) students. Interestingly, however, these closures result in departure of only 296 students, approximately 25 percent fewer departures than under the quality objective. More generally, the contrast of outcomes under the Similarity and Retention objectives illustrates that decisions about school closures can have a large impact on incentives of different groups of students to leave the district.

Minimizing dislocation (Column 4) leads to closure of the same set of schools being closed as under the Quality criterion. This result shows that different objective functions
can yield similar school configuration patterns.

Columns (5) through (8) report the results for the exercise when we also impose the school level capacity constraints. Comparing these results to the ones reported in columns (1) through (4), we find that the results are qualitatively similar. However, there are some important quantitative differences. First note, that imposing the school-level capacity constraints implies that the number of students leaving the district in three out of four scenarios. This is not surprising. The more popular schools are operating much closer to the capacity constraints than the less popular schools. Imposing the school-level constraints then implies that some of the more popular schools will be oversubscribed and hit the capacity constraints. Not having access to their preferred school, a greater number of students decide to leave the district. As a consequence, retaining students is more challenging in the face of school-level capacity constraints.

Results under the similarity objective yield an interesting exception to this pattern. When capacity constraints are taken into account, schools selected for closure serve fewer students and have higher proportions of disadvantaged students. Students induced to leave the district are also more disadvantaged than in the absence of capacity constraints. We also find (not shown in the tables) that the standard deviation of peer characteristics across schools is slightly lower for the capacity versus no-capacity case. Together, these observations indicate that the capacity constraints are limiting the change in peer characteristics in more popular schools and providing less attractive alternatives for students from closed schools. This case illustrates the subtle challenges confronting a district superintendent. Pursuit of an egalitarian objective, expressed in terms of the distribution across schools of students who remain in the district, may result in impacts (e.g., on departing disadvantaged students) that run counter to that objective.

In summary, we find that, regardless of the criterion chosen, closing schools causes significant displacement of students as well as exit of students to outside options. We
also find, however, that the choice of objective function has important consequences for the number and demographic composition of students who are displaced, and for the number and demographic composition of students who exit the district. As urban districts struggle with declining enrollments, they seek policies to retain students while also attempting to avoid having a particular demographic group shoulder a disproportionate share of the adverse impacts of policies. Our computational highlight the tradeoffs among policies and demonstrate the importance of parental school choice decisions to school closing policies. Our results also show the potential of our framework in aiding district management of school capacity decisions.

The state of the art in computational methods limits the scale of school closing problem that we are able to consider. Ongoing research that offers promise of methods for tackling larger problems in the future. It is worth noting, however, that our district confronted a large-scale school closing problem because of years of neglect of declining enrollments and rising excess capacity. The same is true of some other districts now confronting large-scale closing problems. When finally forced to take on large-scale closings, the experience of our district suggests that districts may be more likely thereafter to undertake smaller scale closings as excess capacity emerges. Districts that address closing problems on a more timely basis are likely to confront problems that are closer to a scale of computationally tractable with the approach we have developed.

7 Conclusions

Many large urban districts face declining enrollment due to competition from charter schools and suburban schools and are therefore forced to downsize. To address this important policy issue, we have formulated a sequential game of managing a school district’s student capacity. We have shown that a Perfect Bayesian Equilibrium exists and
characterized its properties. Using data for a mid-sized urban district, we have estimated the parameters of our model. We have shown that consideration of student sorting is vital to the assessment of school closing policy. We find that the district can reduce excess school capacity by closing underperforming schools. This approach inevitably leads to displacement of students in closed schools and additional retention concerns. We have also shown that superintendents confront a difficult dilemma: pursuing an equity objective, such as limiting demographic stratification across schools, results in the exit of many more students than are lost by an objective such as closing underperforming schools. More generally, our results show the feasibility and value of bringing a model of school choice to bear in informing school closing decisions.

An interesting extension of our paper is to include teacher reassignments into the optimal school closing analysis. Downsizing does not only imply that students have fewer options in the public school system. Teachers contracts typically mandate that teachers who were employed in closed schools be retained and be reassigned on the basis of seniority among remaining schools. As a consequence, school quality in the remaining schools will change for two reasons, a change of student peer effects and a change in teacher quality. We do not have access to reliable measures of teacher quality in our analysis. Differences in teacher and principal quality are captured in our estimation by school specific fixed effects which are allowed to change over time. Solving the optimal school closing allowing for endogenous teacher reassignment is exceedingly difficult. However, the optimal school closing analysis problem can be solved with an exogenous policy rule for reassigning teachers. Such an exogenous rule might capture the seniority-driven nature of teacher reassignments. It is then not difficult to extend our framework to incorporate teacher reassignments. These are important avenues for ongoing research to provide districts with more sophisticated tools for addressing problems of declining enrollments that can be expected to stretch well into the future for many urban school
districts. Such tools also have the potential to be extended to help districts with growing enrollments select among potential locations for opening new public schools.

References


Figure 3: The Distribution of the School Ability Coefficient

Ability Coef Dist Middle 90, All Years

A Figures
Figure 4: Distribution of School Black Coefficient

Black Coef Dist Middle 90, All Years

![Distribution of School Black Coefficient](image)

- Black Coefficient
- Frequency

<table>
<thead>
<tr>
<th>Black Coefficient</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2000</td>
</tr>
<tr>
<td>2000</td>
<td>4000</td>
</tr>
<tr>
<td>4000</td>
<td>6000</td>
</tr>
</tbody>
</table>

48
Figure 5: Distribution of School FRL Coefficient
Figure 6: The Distribution of the School Suspensions Coefficient

**susp_days Coef Dist Middle 90, All Years**