Quantitative Economics 9 (2018), 483–520 1759-7331/20180483

The superintendent's dilemma: Managing school district capacity as parents vote with their feet

Dennis Epple
Tepper School of Business, Carnegie Mellon University and NBER

Akshaya Jha
Heinz College, Carnegie Mellon University

Holger Sieg
Department of Economics, University of Pennsylvania and NBER

Many urban school districts in the United States and OECD countries confront the necessity of closing schools due to declining enrollments. To address this important policy question, we formulate a sequential game where a superintendent is tasked with closing down a certain percentage of student capacity; parents respond to these school closings by sorting into the remaining schools. We estimate parents' preferences for each school in their choice set using 4 years of student-level data from a mid-sized district with declining enrollments. We show that consideration of student sorting is vital to the assessment of any school closing policy. We next consider a superintendent tasked with closing excess school capacity, recognizing that students will sort into the remaining schools. Some students will inevitably respond to school closings by exiting the public school system; it is especially difficult to retain higher achieving students when closing public schools. We find 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 explicitly based on student retention.

Keywords. School closing, school choice, demand for public schools, peer effects.

JEL classification. C35, C52, C60, I20.

Dennis Epple: epple@andrew.cmu.edu
Akshaya Jha: akshayaj@andrew.cmu.edu
Holger Sieg: holgers@econ.upenn.edu

We would like to thank the editor of the journal, three anonymous referees, 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).

https://doi.org/10.3982/QE592

1. Introduction

A growing proportion of parents in the United States are moving from cities to suburbs, especially on the East Coast and the Midwest. Due to this, urban school districts in U.S. are increasingly concerned with retaining students. When policies aimed at student retention are not successful, these urban school districts are forced to downsize.[1] Also, OECD countries have experienced low birth rates during the past decades. Sustained low fertility rates will naturally result in less children that need to attend public schools. As a consequence, policymakers in these countries will similarly have to reduce public school capacity.

The objective of this paper is to formulate and estimate a new quantitative framework for evaluating the impact and effectiveness of different downsizing policies.[2] Our framework is based on an extensive form game 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 from the remaining public schools in the district. Parents also have outside options: they can send their children to a private school within the school district or leave the school district altogether. Within our framework, parents choose which school their child attends treating each school as a differentiated product that can be characterized by a vector of endogenous peer measures (such as school-level average test scores) and other exogenous characteristics (such as the driving time from their home to each school). Each parent also has a vector of idiosyncratic shocks (one for each school) that is private information.

We show that a Nash equilibrium of our extensive form game can be characterized using a standard backward induction argument. Given a set of open schools in the second stage, parents have beliefs about the peer qualities of each school and choose the optimal school for their children. Equilibrium in this second stage subgame requires that parents' beliefs about each school's endogenous peer characteristics are consistent with equilibrium sorting. We show that at least one sorting equilibrium exists for any fixed set of schools ("existence"); however, it is theoretically possible that there are multiple sorting equilibria for a given set of schools ("nonuniqueness").

In this first stage of our extensive form game, the superintendent maximizes his or her objective function subject to the parental sorting in the second stage of the game, school-specific student enrollment capacity constraints, and a constraint that limits the overall level of district-wide public school capacity. We show that this optimization problem has at least one solution. However, the fact that the parental sorting stage of the game potentially has multiple equilibria implies that the full two-stage game may have

[1] Between 2001 and 2009, Chicago closed 44 schools; Chicago closed an additional 50 schools in 2013. Detroit closed more than 100 schools over the last decade. The School District of Philadelphia closed six schools in 2012, and closed 24 schools in 2013. Kansas City, Milwaukee, Pittsburgh, and Washington closed between 20 and 30 schools each in recent years. 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. For example, see Engberg, Gill, Zamora, and Zimmer (2012).

multiple equilibria as well. To address this multiplicity concern, we solve for a second-stage sorting equilibrium using different starting values for school-level endogenous peer characteristics; we show that the resulting sorting equilibria are not substantially different across these different starting values. Summarizing, it is possible to optimally select which schools to close under various objectives accounting for the fact that students will self-sort into the remaining open schools. Both our quantitative framework and our empirical analysis using this framework highlight key issues facing superintendents in practice as they undertake school closings.

To implement our quantitative framework, we use student-level data for school years 2004–05 through 2007–08 (four school years) provided by an urban school district (which prefers to remain anonymous). Our empirical analysis exploits the fact that this district implemented a plan to close a substantial number of schools across the district following the 2005/06 school year. The number of elementary and middle schools declined by approximately 25% after the 2005/06 school year as a result of these closures. Another key advantage of our data set is that the district provides transportation for all students residing within its borders. As a consequence, we also have data from students living in the district that attend private and charter schools. Finally, the panel structure of our data provides us with information on which students enter and leave the school district each year. Summarizing, our data allow us to precisely estimate each parent's preferences for the outside options available to them as well as how these preferences change with different public school closing scenarios.

We estimate the parameters characterizing parents' preferences for schools using a two-step estimator that controls for omitted school characteristics.[3] In the first step, we estimate a panel-level random coefficients logit demand model which allows for school-year fixed effects. Our first-stage estimates of the school-year fixed effects capture the average quality of each school in each year. The second stage of our estimator decomposes these school-year specific fixed effects into their observed and unobserved components. In particular, parents choose schools based both on peer characteristics observed to the econometrician (such as school-level average test scores) as well as unobserved school characteristics (such as teaching quality). However, we cannot simply regress our estimated school-year fixed effects on the observed peer characteristics for each school in each year, as these peer characteristics are the endogenous outcomes of the student sorting process.

We instead apply an instrumental variables estimator to account for the endogeneity of school-level peer characteristics. Our first set of instruments is based on the peer effects constructed using the district's assignment of each student's default school; importantly, students are free to attend any public school within the district regardless of their default assigned school.[4] The school assignment rules used by our district are pri-

[3]Our approach to demand estimation follows standard practices in the literature on differentiated products. Berry (1994) and Berry, Levinsohn, and Pakes (1995) introduced the basic random coefficient model with unobserved product characteristics. Berry, Levinsohn, and Pakes (2004) and Bayer, Ferreira, and McMillan (2007) combine individual-level and market-level data to estimate demand models for differentiated products.

[4]Our set of instruments is similar to those constructed by Hoxby and Weingarth (2006) which exploits the reassignment process used by Wake County.

marily a function of distance; the district tends to assign a student to the school that is closest to where they live. Thus, these predicted compositions of schools primarily capture historical, spatial residential sorting patterns.[5] Our identification strategy implicitly relies on the fact that adjustments in residential housing markets are slower and more costly than adjustments of school choices. For example, a parent may choose to reside in a given neighborhood in order to be close to a school with favorable unobserved school characteristics such as better teachers, better facilities, etc.; if these unobserved school characteristics unexpectedly become worse over time, it is easier for this parent to simply send their child to a different public school rather than change residences. As a test of this identifying assumption, we also run our instrumental variable regression only for the school year 2006/07 immediately after our district closed down a large number of schools; these school closings were largely unanticipated by parents in the district, making it very unlikely that parents could change residences based on changes in unobserved school characteristics. Our second stage empirical results are similar whether we restrict to the 2006/07 school year or consider our full 2004/05–2007/08 sample, providing evidence that unobserved school characteristics are not driving our estimates of parents' preferences for observed peer characteristics.[6]

Our empirical estimates provide evidence in favor of the similarity hypothesis: students sort into schools with a higher proportion of other students similar to them. Concretely, we find that parents exhibit preferences for schools with a higher proportion of students of their race, a higher proportion of students with their free or reduced lunch status,[7] and similar student achievement as measured by test scores. This sorting on similarity is crucial for understanding the types of objectives that are difficult for the superintendent to implement; for example, it is difficult given this sorting on similarity to close schools in order to increase student diversity at each of the remaining schools.

We also find that the costs associated with switching schools are relatively large; this highlights the importance of our panel data on parents' choices of schools for estimating their preferences over time. Finally, we show that parents exhibit a marked preference for schools near where they live (as measured by driving times from each school to their home). Both of these results regarding switching costs and driving times work in favor of the superintendent retaining students after closing schools; if the superintendent's only objective is to retain students at public schools, he or she should focus on closing schools with students that are especially immobile (as measured by these switching costs and moving costs).

Finally, we formally specify the optimization problem of a superintendent tasked with closing down 5% of his or her school capacity, subject to the constraints implied by

[5]In this sense, our identification strategy has similarities to the spatial regression discontinuity design advocated by Black (1999).

[6]Our empirical results are also similar if we use lagged peer characteristics as instruments rather than default school assignments, as the sorting equilibrium in the previous period should not be based on unobserved school characteristics from the current period. This identification argument assumes that unobserved school characteristics are not autocorrelated.

[7]Some students receive either free lunch at school or subsidized lunch at school; if a student receives a free or reduced cost lunch, this student on average prefers schools with a higher proportion of students that also receive free or reduced cost lunch.

students self-sorting into the remaining schools and the enrollment capacity constraints of the remaining public schools. We utilize a nested fixed-point algorithm to solve this optimization problem. In the inner loop, we solve for a sorting equilibrium for a given set of school closings, using the parental preferences estimated from our demand model. In the outer loop, we iterate over the set of potential school closing options, picking the option that maximizes the superintendent's objective function.[8] We consider three potential objective functions: (1) closing down the school(s) with the least number of students, (2) closing down the school(s) that result in the least number of students leaving the public school system, and (3) closing down the school(s) that result in the lowest standard deviation across school-level peer characteristics.

The school district in actuality closed down a substantial number of "underperforming" schools following the 2005/06 school year, where the school district identified underperforming schools based on estimated school fixed effects from student-level test score regressions. We can thus recreate the school district's actual objective function. However, the district did not explicitly consider the resorting of students when closing public schools; our optimal school closing algorithm prescribes closing a different set of schools than those actually closed for all three of the objective functions we consider. One mechanism underlying this difference in the set of schools closed is that our estimated unobserved school characteristics are only weakly correlated with the quality measure based on student test scores used by the district. This highlights the importance of considering parental perceptions of school quality in school closing decisions, since students will sort into the remaining open schools based on these perceived school qualities.

Our school closing counterfactuals indicate that closing down schools inevitably results in the exit of students from the public school system; moreover, the students that exit tend to be higher-achieving, which works against closing schools based on an achievement-based quality metric. This finding again highlights the importance of student sorting, as school quality is based in part on peer characteristics that can change drastically for different school closing scenarios. Our results also suggest that 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. The difficulty in retaining students when using a "diversity" based criterion is due to students' preferences to "sort into similarity" in response to school closings.

The rest of the paper is organized as follows. Section 2 discusses how our paper relates to the existing literature. Section 3 describes the data used in the empirical analysis. Section 4 develops our sequential game of managing public school capacity and student sorting for a given set of schools. We specify the estimator for the parameters of our parental demand model for schools in Section 5; Section 6 provides these parameter estimates. Section 7 presents our main findings from quantitative simulations of various school closing scenarios. Section 8 offers conclusions and discusses future research.

[8]Alternatively, we can formulate our optimal school closing problem as a Mathematical Program with Equilibrium Constraints (MPEC). This MPEC alternative to the nested fixed-point algorithm is discussed in Su and Judd (2012) and Su (2014).

2. Literature review

Our paper is related to a growing body of empirical work on school choice. Neal (1997), Altonji, Elder, and Taber (2005), and Card, Dooley, and Payne (2010) analyze various performance metrics for Catholic schools. There has also been substantial research on the effects of school vouchers in Milwaukee (Rouse (1998), Chakrabarti (2008), and Witte, Carlson, Fleming, and Wolf (2012)), Florida (Figlio and Rouse (2006)), Chile (Hsieh and Urquiola (2006), Bravo, Mukhopadhyay, and Todd (2010), Neilson (2013)), and Columbia (Angrist, Bettinger, Bloom, King, and Kremer (2002) and Angrist, Bettinger, and Kremer (2006)). Cullen, Jacob, and Levitt (2006) studied public school choice and student achievement using high school lotteries in magnet programs in Chicago, Hoxby, Murarka, and Kang (2009) studied charter schools in New York, and Abdulkadiroglu, Angrist, Dynarski, Payne, and Pathak (2011) studied charter and pilot schools in Boston. In contrast to those papers, our school quality measures are not based on value-added achievement regressions, but instead reflect parental perceptions of school quality as revealed by observed school choices. Our approach to demand side estimation most closely resembles work by Hastings, Kane, and Staiger (2009); Hastings, Kane, and Staiger (2009) studied the link between parental preferences and student achievement within a random-coefficients logit framework using parental rankings of schools under the Charlotte–Mecklenburg controlled choice system. They found that choice may widen rather than narrow gaps in achievement across demographic groups. Similarly, Neilson (2013) estimated 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.[9] We focus on managing the school district's student capacity using a new game theoretic formulation of equilibrium in the market for public education.

Our approach is also related to a more recent literature that estimates equilibrium models of educational markets. Epple and Sieg (1999) considered sorting of households in a metropolitan market and showed that household stratification by income is primarily driven by difference in school spending across districts. Ferreyra (2007) estimated 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 papers approach school choice as a mechanism design problem. We formulate a game in which the superintendent is tasked with closing down school capacity facing similar individual rationality and self-selection constraints based on parents' school choices.

3. Data

Similar to many urban school districts, the district that we study in this paper has been experiencing declining public school enrollment. The fiscal pressures associated with

[9] Recent work by Dinerstein and Smith (2015) examines the effect of public school funding on private school closure decisions in New York City.

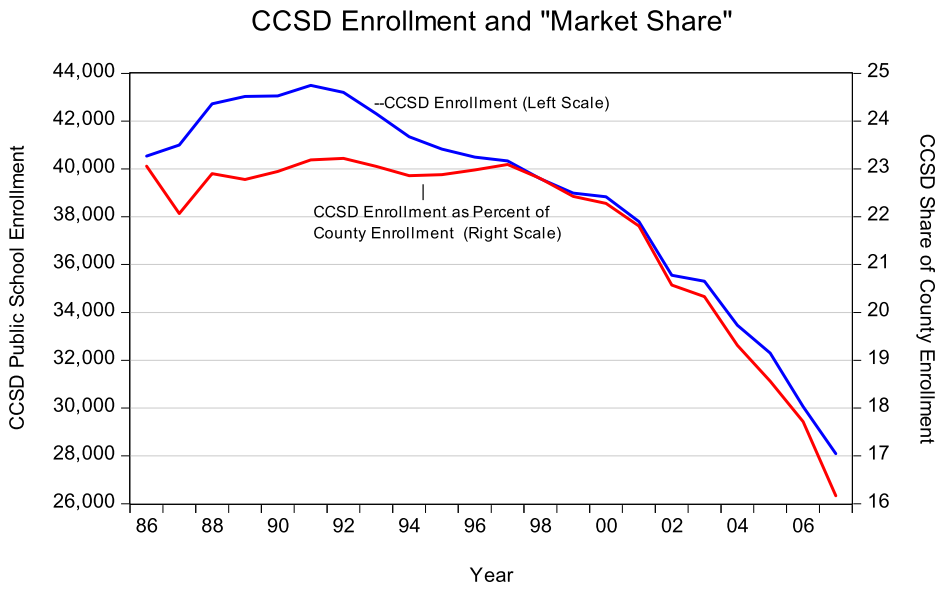

FIGURE 1. Enrollment and market share: 1987-2006.

this declining enrollment have necessitated closing of schools. We refer to our district for the rest of the paper as the Center City School District (CCSD) since the district prefers to remain anonymous. CCSD is located in a county that contains more than 40 suburban school districts and is home to approximately 60% 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 both the yearly number of students enrolled in public schools located within CCSD as well as the yearly proportion of students in all districts in the county that were enrolled in public schools located within in CCSD (CCSD's student "market share"). CCSD maintained its student market share during the 1990s when enrollment in the district's schools was rising, but its market share dropped rapidly when metropolitan enrollment began to decline. Countywide births started to decline in the early 1990s; much of this decline in births was a result of the end of the "Echo Boom." Countywide enrollment began to decline in 1998 largely as a result of this decline in births beginning in the early 1990s. CCSD experienced a disproportionate decline of the decline in countywide student enrollment, as evidenced by its decreasing student share of county enrollment.

To explore this phenomenon, we rank school districts in the county by income and aggregate them into quartiles of roughly equal enrollment. We find that more affluent districts did not lose many students 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"); CCSD bore 75% of the countywide decline in public school enrollment as a consequence. This decline in public school enrollment resulted in many CCSD public schools underutilizing their student capacity, which placed significant fiscal strain on CCSD as education funding

TABLE 1. Summary statistics of the CCSD sample K–8.

	2004	2005	2006	2007
Sample size	24,660	24,876	24,489	23,735
Sample size: public + private	22,158	21,189	20,333	19,045
Sample size: public	21,239	20,180	19,189	17,956
Free or reduced lunch	17,438	17,840	17,446	16,866
Race: black	14,256	14,219	13,866	13,312
Race: white	8498	8545	8405	8164
Race: other	1906	2112	2218	2259
Moving cost indicator	10,938	9920	12,255	8796
Driving times (in minutes) to school attended	3.086	3.130	3.273	3.362
Individual achievement measure	−0.042	−0.033	−0.031	−0.045

is primarily enrollment-based and there are sizable fixed costs associated with keeping a school open. In 2005, CCSD launched a series of initiatives aimed at addressing these challenges. This led to the closure of 22 schools serving students in kindergarten through eighth grade at the end of the 2005–06 school year.

Our empirical analysis focuses on the impact of school closings on parental choices in CCSD. Our sample consists of all K–8 students in CCSD that attended public schools for at least one year in 2004–2007; we consider 4 years of data spanning the 2004/05 school year to the 2007/08 school year. We exclude all private school students that never attended a public school from our sample for two reasons. First, we find in our data that very few students that attend private schools ever return to public schools. Second, we do not observe standardized test scores for students that never attended public schools, which is a crucial input into our empirical model of parental demand for schools. We 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. We define the variables used in this study as follows. For the moving indicator, we count as "moved" any student who attended a different school than in the previous year. These costs of moving to a new school capture factors such as having to acclimate to new facilities, teachers, and peers. Our categorization includes students moving into the district, and students changing from elementary to middle school. We do this because students moving from elementary to middle school face similar sorts of moving costs as those switching schools in another grade. The driving times variable denotes the median driving time from home to school in minutes. Student achievement is measured as an average of all observed standardized test scores; we first standardize each of these test scores (e.g., English, Math, etc.) to have mean zero and a standard deviation of one before taking the average for each student in each year.

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

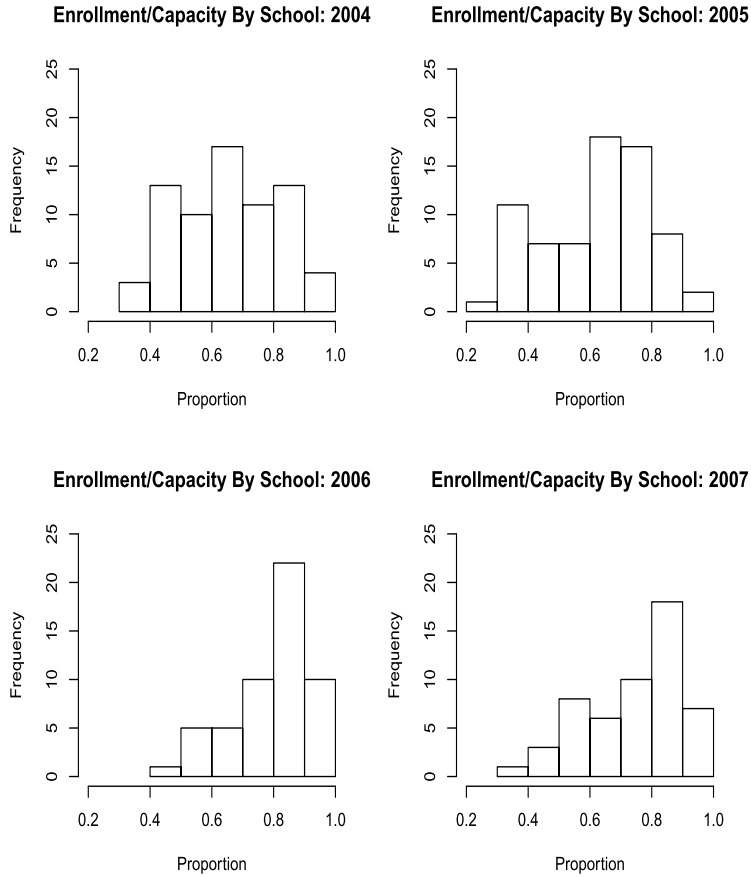

FIGURE 2. School capacity: 2004-2007

school during the years 2002–2007; we denote this magnitude the "observed capacity" of a school. Then we simply take the maximum of stated and observed capacity to create the capacity measure necessary for our optimal closing analysis.[10] We find that the district had an overall capacity of 34,053 K–8 students in 2005 before the school closing. The capacity fell to 24,588 students after the school closings were implemented after the 2005/06 school year. For grades K–5 (6–8), the capacity was reduced from 13,192 (20,861) to 9037 (15,550). Figure 2 plots the empirical distribution of school-level capacity utilization—the fraction of student enrollment in a school divided by their student capacity—for all of the schools in the district. We find that only a few schools operated near capacity before the reforms; there was still some excess capacity left in the CCSD system after the reforms.

Finally, 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

[10]We also 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.

student achievement," and is based on a combination of regression specifications involving student test scores as the dependent variable.[11] The SPI is a categorical ranking from the set $\{1, 2, 3, 4\}$, with 1 being the worst possible ranking and 4 being the best. CCSD closed 22 schools (primarily those with SPI $= 1$) using this categorical measure. We convert these categorical variables into continuous variables by regressing them on the average achievement measures reported by the district in order to implement our counterfactual school closing analysis. We consider the predicted SPI as the school quality ranking implied by the CCSD's objective function in our optimal school closing analysis.

4. Managing school district capacity

We consider a sequential game in extensive form played by the superintendent of a school district and a continuum of parents. The superintendent determines the set of schools that are open in the first stage of the game. In the second stage, parents enroll their children in one of the (open) public schools or choose one of the outside options (consisting of a charter school, a set of private schools, or leaving the school district). Parents choose schools taking into consideration that the peer characteristics of these schools are the endogenous outcomes of a sorting equilibrium. We begin by discussing the second stage of the game where students sort into schools and then formulate the optimization problem facing a superintendent tasked with closing school capacity given this second stage sorting equilibrium. We show that a Nash equilibrium to our sequential game exists using a standard backward induction argument.

4.1 *The second stage: School choice with peer effects*

Let J_t denote the set of potential public schools that are available in school year t.[12] We denote the set of schools that are to be closed $J_t^C \subseteq J_t$, while we call the set of public schools that remain open J_t^O. By construction, we have $J_t = J_t^O \cup J_t^C$.

We make two assumptions consistent with our empirical setting regarding the 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 assume when estimating our parental demand model for schools that parents observed school choices are not based on binding school-level capacity constraints; each parent chooses their most preferred school for his or her child.[13] However, we do account for school-level capacity constraints in our counterfactual school closing analysis; namely, if the superintendent decides to counterfactually

[11] Neilson (2013) used a similar approach to measure school quality.

[12] We account in our empirical analysis for the fact that the number of school options available to a student depends on their grade. However, we suppress this dependence on grade in the exposition for notational convenience.

[13] From Figure 2, we see that almost all public schools in the district have excess capacity. However, there was a small number of selective magnet programs that were operating at capacity and used lotteries to determine admissions during our sample time period. Our empirical results are similar if we include or exclude students attending oversubscribed magnet programs.

close down a set of schools J_t^C, we let students sort into the remaining schools subject to school-level capacity constraints. We discuss how we ration this public school capacity in our counterfactual analysis in the next subsection.

We denote the set of outside options O_t. The choice set is thus given by $J_t^O \cup O_t$; we assume in our empirical application that each student can attend a (generic) charter school, three different parochial school types (e.g., Catholic private school), an independent private school within the district, or be home-schooled. In addition, the student can choose to attend a school outside of the district. We do not model heterogeneity in schools within these outside option types; for example, we do not model differences in school quality across different charter schools within the district. Parents exercise school choice by enrolling their children in one of the district's public schools or one of the outside options.

We treat each school j as a differentiated product with a combination of endogenous characteristics ($\bar{z}_{j,t}$) and exogenous characteristics ($\xi_{j,t}$). Endogenous characteristics depend on the outcome of the sorting process. In our application, $\bar{z}_{jt}$ includes peer characteristics such as the average achievement of students in each school and the average number of student suspensions in each school as well as demographic variables such as the proportion of students eligible for subsidized lunch and the proportion of students in different racial/ethnic groups. An example of exogenous characteristics include the quality of the principal and the teachers. When parents make decisions about enrolling their children into public schools, they hold beliefs about each school's peer effects; these beliefs have to be consistent with the optimal strategies of parents and the superintendent in equilibrium.

We let $z_{i,t}$ be the observed vector of characteristics of student i for school year t. Variables in $z_{i,t}$ include the student's achievement, race, free or reduced lunch status, as well as number of suspensions. Let $d_{i,j,t-1}$ be an indicator variable which is equal to one if student i attended school j in the previous school year $t-1$. Previous school choices matter in our model because transferring to a new school is costly; namely, switching schools requires the student to adapt to new school rules and acclimate to new peers, facilities, and teachers. Finally, we also include the driving time from the student's home to each school in her choice set; we suppress this variable in the discussion of our model for expositional simplicity.

Each parent i has a preference shock for each school j in each school year t; we denote the parent's vector of shocks for each school $\varepsilon_{i,t}$. We assume that parents' preference shocks are private information. Each parent also has a vector of preferences over endogenous school characteristics, which we denote β_i. A student is therefore completely characterized by her vector of characteristics $(z_{i,t}, d_{i,t-1}, \beta_i, \varepsilon_{i,t})$.[14] We assume that the utility function of student i in school year t is additively separable in her idiosyncratic preference shocks and can thus be written:

$$U(\bar{z}_t, \xi_t, z_{i,t}, d_{i,t-1}, \beta_i, \varepsilon_{i,t}) = \sum_{j \in J_t^O \cup O_t} d_{i,j,t}\big[u(\bar{z}_{j,t}, \xi_{j,t}, z_{i,t}, d_{i,t-1}, \beta_i) + \varepsilon_{i,j,t}\big]. \tag{1}$$

[14]We consider the parent–student pair as a single decision-making agent, and will thus use "parent" and "student" interchangeably.

Parents choose the school that maximizes their utility given their set of available school options and their beliefs about endogenous school characteristics; each parent employs a pure strategy as they choose only one school for their child. However, each parent's decision is random from the perspective of the superintendent and other parents due to the existence of private information. Integrating out the private information yields conditional choice probabilities for each student i for each school option j in school year t:

$$\Pr\big[d_{i,j,t} = 1 | J_t^O, \bar{z}_t, \xi_t, z_{i,t}, d_{t-1}\big]. \tag{2}$$

Summing these probabilities across all parents yield the aggregate market shares for each public school and each outside option in each school year; we denote these market shares $s_{j,t}$.

These conditional probabilities result in an equilibrium of the second stage of the game if school-level peer characteristic $\bar{z}_{j,k,t}$ satisfies the following consistency requirement for all schools j, all school years t, and all peer characteristics k (i.e., student achievement, free and reduced lunch status, race, and number of suspensions):

$$\overline{z}_{j,k,t} = \frac{\sum_{i=1}^{N_t} z_{i,k,t} \Pr\big[d_{i,j,t} = 1 | J_t^O, \bar{z}_t, \xi_t, z_{i,t}, d_{t-1}\big]}{\sum_{i=1}^{N_t} \Pr\big[d_{i,j,t} = 1 | J_t^O, \bar{z}_t, \xi_t, z, d_{t-1}\big]} \quad \forall (j,k,t), \tag{3}$$

where N_t is the number of students in school year t. In words, students sorting according to the conditional probabilities implied by equation (2) must result in the school-level peer characteristics that were used to generate these probabilities. Parents' beliefs regarding school-level peer characteristics are confirmed in equilibrium.

We further assume that each pubic school has a capacity constraint equal to $K_{j,t}$. Equilibrium in the second stage of the game requires that each public school's capacity constraints are satisfied:

$$\sum_{i=1}^{N_t} \Pr\big[d_{i,j,t} = 1 | J_t^O, \bar{z}_t, \xi_t, z_{i,t}, d_{t-1}\big] \leq K_{jt} \quad \text{for all } j \in J^O. \tag{4}$$

We consider these capacity constraints when computing an equilibrium of the second stage of the game. In particular, we use a rationing rule that is based on the idea that parents are charged a shadow price for admission given by $p_{j,t}$. Parents who want their children to attend a school facing excess demand have to invest extra 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. The reduction in demand occurs because the shadow price is deducted from the utility that parents obtain from having their child attend the school.[15] We normalize

[15]Some urban districts ration overdemanded schools by lottery. A different equilibrium analysis and estimation strategy would be required for such scenarios. Geyer and Sieg (2013) modeled and estimated the demand for public housing. They showed how to modify the demand system in the presence of lotteries and compute the endogenous probabilities that determine the outcome of the rationing process.

the shadow price of leaving the district to be zero. It is straightforward to show that this rationing device is efficient; it allocates the available capacity to those who value it most.

Formally, we measure the shadow price in the same units as the unobserved school characteristic; it enters additively into the utility function. Each parent's utility function takes the form:

$$U(\bar{z}_t, \xi_t - p_t, z_{i,t}, d_{i,t-1}, \beta_i, \varepsilon_{i,t}) = \sum_{j \in J_t \cup O_t} d_{i,j,t}\big[u(\bar{z}_{j,t}, \xi_{j,t} - p_{j,t}, z_{i,t}, d_{i,t-1}, \beta_i) + \varepsilon_{i,j,t}\big].$$

Summarizing, we characterize a sorting equilibrium for a given set of schools $J_t^O \cup O_t$ based on the following two conditions:

1. All peer characteristics k for all schools j ($\bar{z}_{j,k,t}$) are consistent with a student sorting equilibrium:

$$\bar{z}_{j,k,t} = \frac{\sum_{i=1}^{N_t} z_{i,k,t} \Pr\big[d_{i,j,t} = 1 | J_t^O, \bar{z}_t, \xi_t^O - p_t^O, z_{i,t}, d_{t-1}\big]}{\sum_{i=1}^{N_t} \Pr\big[d_{i,j,t} = 1 | J_t^O, \bar{z}_t, \xi_t^O - p_t^O, z, d_{t-1}\big]} \quad \forall (j,k,t).$$

2. The market share of each school $s_{j,t}$ is less than or equal to the capacity constraint ($s_{j,t} \leq K_{j,t}\ \forall (j,t)$).

4.2 *The first stage: Optimal configurations*

In the first stage of our sequential game, a superintendent is tasked with closing down a certain percentage of public school capacity with the knowledge that students will then sort into the remaining schools. The superintendent may have a number of objectives in deciding which schools to close down. We first characterize the objective function actually used by the district in practice. We then discuss alternative objectives that districts in general may have in mind when closing down public schools.

Superintendents often manage capacity with the explicit goal of closing "underperforming" schools. As we discussed in Section 3, CCSD estimated school-specific, achievement-based quality measures ($q_{j,t}$) using student-level test score data which they used when closing down schools following the 2005/06 school year. These quality measures were treated by the district as exogenous attributes of the schools (i.e., CCSD implicitly assumed that they did not depend on sorting by households). A good approximation of the school district's decision problem is that the district maximizes a weighted average of school quality:

$$Q_t = \sum_{j \in J^O} w_{j,t} q_{j,t}, \tag{5}$$

where $w_{j,t}$ denotes the weight of school j in school year t.

The objective function in equation (5) provides no role for considerations other than school performance that are potentially important to the 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 equation (5) ignores the dependence of choices on students' proximity to schools. This may make it difficult to retain students, especially if some students are forced to commute longer distances. Finally, maximizing equation (5) may also imply large reallocations of students among schools, causing a large number of students to incur the sizable costs associated with switching schools. Of course, the district was almost certainly cognizant of these potential ramifications of school closings in pursuing the objective in equation (5); CCSD 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.

We consider three alternative objectives in order to gain insight into these important issues: limiting demographic stratification across schools, retaining students in the district, and minimizing the number of students attending schools chosen for closure. We show that these different objective functions lead to different schools being closed as well as vastly different school-level peer characteristics. Thus, the superintendent in practice faces a difficult dilemma in balancing these different objectives.

We use the sum of the weighted squared deviation between school j's and the district's characteristics over public schools to measure inequality in the provision of education. Let $\overline{z}_{k,t}^{J^O}$ denote the average of peer characteristic k over all open public schools in the district for school year t. We find the standard deviation of each peer characteristic k over all open public schools J_t^O, and then take the weighted sum over peer characteristics in order to construct our total inequality index:

$$I_t = \sum_{k=1}^{K} \sum_{j \in J^O} \omega_{k,t} \big(\overline{z}_{j,k,t} - \overline{z}_{k,t}^{J^O}\big)^2, \tag{6}$$

where $\omega_{k,t}$ is the weight assigned to school characteristic k. For our computational analysis, we use numerical values for the weights ($\omega_{k,t}$'s) that result in each peer characteristic having approximately equal weight in the objective function.[16] The superintendent wants to minimize the weighted sum of the standard deviations of each school-level peer characteristic in order to increase diversity.

The district also wants to attract and retain students in their public schools; thus, the superintendent wants to minimize the number of students leaving for outside options (such as private schools or leaving the district):

$$R_t = \sum_{i=1}^{N_t} \sum_{j \notin J^O} \Pr\big[d_{i,j,t} = 1 | J_t^O, \bar{z}_t, \xi_t^O, z_{i,t}, d_{i,t-1}\big]. \tag{7}$$

[16]Namely, we use weights to adjust for the different scales of different peer characteristics; for example, the average number of student suspensions for each school is typically larger than one while the proportion of students on free or reduced lunch is bounded between zero and one.

Finally, the district wishes to limit the number of students that are attending schools that are closed in order to limit the moving costs incurred by students forced to switch schools. The number of students in closed schools is

$$D_t = \sum_{i=1}^{N_t} \sum_{j \in J_t^C} \Pr\left[d_{i,j,t-1} = 1 | J_{t-1}^O, \bar{z}_{t-1}, \xi_{t-1}, z_{i,t}, d_{i,t-2}\right]. \tag{8}$$

Based on these various objectives, the school district is tasked with closing down at least a certain proportion of district-level public school capacity c (we empirically consider $c = 0.05$); formally, it faces the following constraint:

$$\sum_{j \in J^O} K_{j,t} \leq (1 - c) \sum_{j \in (J^O \cup J^C)} K_{j,t}, \tag{9}$$

where $K_{j,t}$ is the capacity of school j. Finally, the superintendent faces school-level capacity constraints, setting shadow prices in order to ration students such that no public school's enrollment is above their capacity.

We can characterize the trade-offs faced by the superintendent of a school district in closing down schools by studying the equilibrium properties of our sequential game for alternative objective functions. We argue that an equilibrium for our sequential game exists using a backward induction argument. First, the second stage of our game is almost identical to the neighborhood sorting game developed in Bayer and Timmins (2005) if we ignore school-specific capacity constraints; Bayer and Timmins (2005) showed that the existence of equilibrium without rationing follows from a straightforward contraction mapping argument. In our setting, we need to iterate on the shadow prices for each school in order to ration any excess demand that arises from the equilibrium sorting of students. We do not prove theoretically that this sorting model with rationing is a contraction mapping; instead, we show that our empirical framework converges to a sorting equilibrium for a wide range of starting values for the school-level peer characteristics. This provides strong evidence that a contraction mapping continues to exist even with rationing; we thus conclude that at least one second-stage equilibrium exists based on the argument provided in Bayer and Timmins (2005) and can be computed for our empirical optimal school closing framework.

As in most models with endogenous peer effects, there is some scope for multiplicity of equilibria. We address this concern by using different starting values for peer characteristics for a given set of open public schools in order to explore whether this multiplicity concern is relevant empirically. We find that the peer characteristics generated from solving for a sorting equilibrium are not substantially different when using different starting values for the peer characteristics, indicating that multiplicity of equilibria is not a concern empirically.

At least one Nash equilibrium for the game in extensive form exists since there are only a finite number of feasible school configurations. Also, the equilibrium to the overall sequential game is unique if and only if the equilibrium to the second stage sorting game is unique for all possible school closing scenarios. We can find an equilibrium for this sequential game for reasonably-sized problems by simply evaluating all possible

combinations of school closings. However, the superintendent's optimization problem increases exponentially in complexity with the number of schools to be closed, making iteration over of all school closing alternatives an infeasible approach for large problems.[17]

4.3 *An algorithm for computing equilibria to this game*

We consider a nested fixed-point algorithm for solving the superintendent's optimal school closing problem.[18] This algorithm consists of an inner loop and an outer loop. The inner loop solves for a sorting equilibrium for a given set of schools that remain open (J^O). This inner loop is an iterative algorithm where each iteration n consists of the following steps:

1. At the beginning of iteration n, we have a vector of school characteristics denoted by $\bar{z}_j^n$ and conditional choice probabilities for each student i and school j ($\mathbf{P}_{i,j}^n$).[19]

2. In this step, we need to compute the new conditional choice probabilities $\mathbf{P}_{i,j}^{n+1}$ for each student i and school j implied by equation (2) evaluated at school characteristics $\bar{z}_j^n$. However, we must first set shadow prices p_j^n in order to actually compute $\mathbf{P}_{i,j}^{n+1}$. We do this by solving the following constrained optimization problem: we minimize over shadow prices p_j^n the squared difference between the new conditional choice probabilities and the baseline probabilities where no school is closed subject to the constraint that all school-level enrollments are less than their capacities.[20] We normalize the shadow price of leaving the school district to be zero ($p_J^n = 0$ for all n).

3. We compute new school characteristics for each school j as the probability-weighted average of student characteristics (z_i) using these conditional choice prob-

[17]Future applications of our framework may be able to draw on ongoing work in operations' research. One branch of this work seeks to enhance the speed and accuracy of methods for computing an approximate solution to mixed-integer nonlinear programming problems (Belotti, Lee, Liberti, Margot, and Waechter (2009)); these methods also provide a measure of the extent by which this 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)).

[18]This nested fixed-point algorithm was first introduced by Rust (1987) in the context of estimating single-agent dynamic models.

[19]For the baseline case where we do not close any schools, we initialize school characteristics $\bar{z}_j^1$ as the actual school characteristics, all shadow prices p_j equal to zero, and $\mathbf{P}_j^1$ as the school choice probabilities estimated from a conditional logit model with all shadow prices equal to zero evaluated at the actual school characteristics. For the case where we close down some subset of public schools, we initialize the conditional probabilities, shadow prices, and school characteristics at the final solutions found for the baseline case. We vary these starting values as a sensitivity analysis.

[20]We solve for a sorting equilibrium using this inner loop algorithm with all public schools open in order to compute baseline probabilities. Our inner loop algorithm for a given set of closed public schools implicitly picks the closest possible sorting equilibrium to the one computed for the baseline case where no public schools are closed. Therefore, differences between the baseline and counterfactual closing scenarios would be even larger if we instead chose a different sorting equilibrium for the counterfactual closing scenario.

abilities $\mathbf{P}_{i,j}^{n+1}$:

$$\bar{z}_j^{n+1} = \frac{\sum_{i=1}^{N} \mathbf{P}_{i,j}^{n+1} z_i}{\sum_{i=1}^{N} \mathbf{P}_{i,j}^{n+1}}.$$

We iterate on these steps until the maximum absolute difference between the school characteristics in iteration m ($\bar{z}_j^m$) and the school characteristics in iteration $m+1$ ($\bar{z}_j^{m+1}$) is below a given tolerance; empirically, we consider $\max_k(\frac{|\bar{z}_{j,k}^m - \bar{z}_{j,k}^{m+1}|}{\text{scale}(z_k)}) < 1\text{e}{-5}$, where j indexes school, k indexes peer characteristics, and $\text{scale}(z_k) \equiv \max_j(z_{j,k}^{\text{actual}}) - \min_j(z_{j,k}^{\text{actual}})$. We use this scaling factor $\text{scale}(z_k)$ in order to standardize differences in magnitude between different school characteristics (e.g., the proportion of free and reduced lunch students must be between zero and one while school-level average achievement ranges from roughly -1 to 1).

In the outer loop, we iterate over all possible public school closing scenarios J^C that close at least 5% of district-wide public school capacity. In particular, for each possible set of schools closed J^C, we solve for a sorting equilibrium using the inner loop algorithm described above and then evaluate each of our four objective functions for the superintendent. The inner loop of this nested fixed-point algorithm implicitly finds the sorting equilibrium that is closest to the one observed in the baseline case where no schools are closed; we explore multiplicity of sorting equilibria in this nested fixed-point algorithm by varying the initial values of the peer characteristics fed into the inner loop of the algorithm.[21] We minimize each of our four objective functions over the different possible combinations of schools to close in the outer loop, solving for a sorting equilibrium for each public school configuration in the inner loop. We show empirically that different objective functions result in different prescriptions regarding which schools to close, highlighting the trade-off between objectives that superintendents face when closing schools.

5. Estimating parental demand for schools

5.1 *Notation and setup*

We estimate parental preferences for schools using standard techniques developed in the differentiated products demand literature. We assume that the utility of student i in school j in school year t is given by

$$U_{i,j,t} = \sum_{k=1}^{K} \bar{z}_{j,k,t} \beta_{i,k,t} + \xi_{j,t} + \varepsilon_{i,j,t}, \tag{10}$$

[21] Alternatively, one can recast our two-stage school capacity game as a Mathematical Programming with Equilibrium Constraints (MPEC) optimization problem, where the superintendent maximizes his or her objective function subject to the three constraints described above. This MPEC formulation implicitly selects the sorting equilibrium for each possible school closing that maximizes the superintendent's objective function.

where the kth characteristic of school j is denoted by $\bar{z}_{j,k,t}$ and

$$\beta_{i,k,t} = \alpha_{0,k} + \sum_{m=1}^{M} \alpha_{1,k,l} z_{i,m,t} + \sigma_k u_{i,k}, \tag{11}$$

where $z_{i,m,t}$ is the mth component of student i's characteristics in school year t. The random coefficient errors are time invariant and satisfy $u_{ik} \sim N(0,1)$.[22] Define the fixed effect of school j in year t as

$$\delta_{j,t} = \sum_{k=1}^{K} \alpha_{0,k} \bar{z}_{j,k,t} + \xi_{j,t}. \tag{12}$$

We can then write the school specific utility of individual i in year t as

$$u_{i,j,t} = \delta_{j,t} + \sum_{k=1}^{K} \sum_{m=1}^{M} \alpha_{1,k,m} \bar{z}_{j,k,t} z_{i,m,t} + \sum_{k=1}^{K} \sigma_k \bar{z}_{j,k,t} u_{i,k} + \varepsilon_{ijt}. \tag{13}$$

We have ignored travel times when writing the school-specific utility above for expositional simplicity. We have also omitted student moving costs which are given by $mc_{i,j,t} = \gamma_{i,t} 1\{d_{i,j,t} \neq d_{i,k,t-1}\}$ where $\gamma_{i,t} = \gamma_0 + \sum_{m=1}^{M} \gamma_{1,l} z_{i,m,t}$. We add these terms to the model specification when we estimate the model.[23]

Idiosyncratic shocks in the utility function, $\varepsilon_{i,j,t}$, follow a Type I extreme value distribution (McFadden (1974)). Thus, parents' probabilities of attending each school conditional on the observed characteristics $(\bar{z}_t, z_{i,t}, d_{i,t-1})$ and the unobserved differences in parents' preferences for each peer characteristic k $(u_{i,k})$ are

$$Q_{i,j,t} = \frac{\exp\left(\delta_{j,t} + \sum_{k=1}^{K} \sum_{l=1}^{L} \alpha_{1,k,l} \bar{z}_{j,k,t} z_{i,l,t} + \sum_{k=1}^{K} \sigma_k \bar{z}_{j,k,t} u_{i,k}\right)}{\sum_{m} \exp\left(\delta_{m,t} + \sum_{k=1}^{K} \sum_{l=1}^{L} \alpha_{1,k,l} \bar{z}_{m,k,t} z_{i,l,t} + \sum_{k=1}^{K} \sigma_k \bar{z}_{m,k,t} u_{i,k}\right)}. \tag{14}$$

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

$$\Pr[d_{i,j,1}, d_{i,k,2}, d_{i,l,3}, d_{i,m,4} | \bar{z}, z_i, u_i, d_{i,n,0}] = Q_{i,j,1} Q_{i,k,2} Q_{i,l,3} Q_{i,m,4}. \tag{15}$$

[22] We also estimated a model that allowed for nonzero correlations among the random coefficients but did not find improvements in model fit. Alternatively, one could generate realistic substitution patterns in demand using a nested logit model as shown by McFadden (1981) and Goldberg (1995).

[23] Moving costs only apply if a student attends a new school that is in a different physical location. For example, there are a small number of K–8 schools in our application; a student that moves from fifth to sixth grade within that school is not considered to be a mover. However, if the student decides to leave that school after fifth grade and attend a different 6–8 middle school, he or she would be considered to be a mover.

However, we need to integrate out the errors to obtain conditional choice probabilities that only depend on observables given that we do not observe u_i:

$$\Pr[d_{i,j,1}, d_{i,k,2}, d_{i,l,3}, d_{i,m,4}|\bar{z}, z_i, d_{i,n,0}] = \int Q_{i,j,1}Q_{i,k,2}Q_{i,l,3}Q_{i,m,4}\,dF(u_i). \tag{16}$$

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

However, our estimation is complicated by the fact that parent i's school choice at time $t = 0$ ($d_{i,0}$) depends on their unobserved, time-invariant tastes for different peer characteristics u_i (the "initial conditions" problem).[24] We address this concern by only using students that have to move no matter which school they choose at time $t = 0$ ($d_{i,j,0} = 1$ for all j); this includes students entering kindergarten, moving to sixth grade from a K–5 school, and students moving into the district.[25] Of course, as all unobservables are assumed to be independent and identically distributed across both parents and school years in the panel conditional logit model, we do not face an initial conditions problem when estimating the conditional logit model without this u_i term.

5.2 *First stage of demand estimation*

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. The conditional choice probabilities in equation (14) depend on the parameters α_1, σ, and the mean utilities $\delta = (\delta_{1,1}, \ldots, \delta_{J,T})$ (where J denotes the number of schools and $T = 4$ denotes the number of school years); we estimate these parameters using a maximum likelihood estimator. Our likelihood function is given by

$$L = \prod_{i=1}^{N} \int Q_{i,j,1}Q_{i,k,2}Q_{i,m,3}Q_{i,s,4}\,dF(u_i). \tag{17}$$

Formally, we identify and estimate the parameters of the likelihood function by appealing to large N (number of students) and finite J (number of schools) asymptotics. We provide an informal discussion below regarding how we identify the variance terms underlying parents' random (from our perspective) tastes for peer characteristics as well as how we address the potential for multiplicity in equilibria in our sorting game (i.e., multiplicity in the ways students can sort into schools).

The intuition behind the identification of the variance terms in our random coefficient logit model in a panel setting is similar to the identification argument used

[24] Wooldridge (2005) discussed this initial conditions problem for dynamic, nonlinear panel data models with unobserved heterogeneity; our panel data random coefficients logit model falls into this more general class of models.

[25] When estimating our model, we include students beginning in year t if they satisfy this condition. For example, if a student moved into the district in 2005/06 ($t = 2$), we would include this student only for school years 2005/06, 2006/07, and 2007/08 (years $t = 2, 3, 4$).

when researchers have data on individuals' second-most preferred choice. For example, consider a student attending a public school with high average achievement that closes down in 2005/06. Suppose for simplicity that this student can now choose to either attend a nearby school with low average achievement or a school with high average achievement that is farther away; the school that the student chooses will help identify whether he has a high or low unobserved taste for peer achievement. These types of observed substitution patterns are not captured by the school-year fixed effects since similar students will make different school choices. In our example, one student may choose to attend the nearby, lower average achievement school while a student with similar observable characteristics may choose to attend the higher average achievement school that is farther away. Summarizing, the observed substitution patterns generated from students changing schools over time identify the variances of the random coefficients on peer characteristics.

Our likelihood function conditions on the average peer characteristics for each school observed in our data. As we discussed in Section 4, these peer characteristics are the result of student sorting based on unobservables and are therefore potentially endogenous. Further, there is scope for multiplicity of equilibria in these types of sorting games; we condition on the equilibrium under which the data were generated by conditioning on the observed peer characteristics for schools. Conditioning on observed equilibrium outcomes also allows us to implement our estimator without having to compute an equilibrium of the sorting model. Finally, we assume that students do not factor in the (negligible) impact they have on the peer characteristics of schools when making their school choice; this assumption is reasonable for our empirical application given the large number of students in CCSD.[26]

5.3 *Second stage of demand estimation*

We estimate school-year specific fixed effects in the first stage of our demand estimation; these school-year fixed effects are a function of observed and unobserved school characteristics:

$$\delta_{j,t} = \alpha_0 \bar{z}_{j,t} + \xi_{j,t}. \tag{18}$$

Our observed school characteristics are: (1) the school-level proportion of Black students, (2) the school-level proportion of students on free or reduced lunch, (3) the school-level average achievement (as measured by test scores), and (4) the school-level average number of student suspensions. These peer characteristics are the outcome of a sorting equilibrium where parents choose schools for their students based both on observed and unobserved school characteristics. Thus, peer characteristics $\bar{z}_{j,t}$ are likely correlated with unobserved school characteristics $\xi_{j,t}$; in short, $E[\xi_{j,t}|\bar{z}_{j,t}]$ is not likely to be zero.

[26] Bayer and Timmins (2005) provided a comprehensive Monte Carlo study which shows that these types of estimators are well behaved in large samples; Bayer, McMillan, and Reuben (2004) also discussed estimating demand in contexts with sorting.

Following Berry (1994), we instead assume that $E[\xi_{j,t}|w_{j,t}] = 0$ for some set of instruments $w_{j,t}$.[27] We construct instruments $w_{j,t}$ using the administrative default school assignment rules implemented by CCSD; we use these default school assignments in order to predict the composition of each school before and after the school closings in 2005/06. The validity of these instruments relies on the assumption that peer measures predicted by administrative assignment rules are not correlated with unobserved school characteristics. We justify this assumption by noting that the default school assignment rules used by our district are primarily a function of distance; the district tends to assign a student to the school that is closest to his residence. Therefore, the predicted compositions of schools can be viewed as a nonlinear transformation of historical, spatial residential sorting patterns. In that sense, our approach to identification has some similarities with the spatial regression discontinuity design proposed by Black (1999).

One can then ask the question what economic and behavioral models of school and residential choice are broadly consistent with our instrumental variable approach. As discussed in Section 3, our school district experienced large changes throughout our sample period; these changes were largely caused by a rapidly shrinking district wide enrollment. Our identification strategy implicitly assumes that parents made residential choices based on the past unobserved school characteristics, but that these unobserved school characteristics underwent significant changes after this residential choice was made.[28] We argue that these changes in unobserved school characteristics for CCSD likely occurred during our sample period, as there were changes in the administration, reassignments of teachers and principals, as well the reconstitution of some of the schools. However, our identification strategy still breaks down if parents made their residential choices fully anticipating these changes in unobserved school characteristics. We argue in Section 3 that many parents and administrators did not fully anticipate the speed and extent of the decline in district-wide student enrollment and its impact on unobserved school characteristics when making their residential choices (for parents) or assigning default schools to each student in the district (for administrators). Finally, our identification strategy relies on the fact that adjustments in residential housing markets are slower and more costly than adjustments of school choices. In particular, we argue that 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, relative to simply moving their children from one public school to another within a district that practices a de facto open enrollment policy. There is a substantial amount of empirical evidence in the literature supporting this assumption.[29] The peer effects predicted

[27] Berry, Linton, and Pakes (2004) discussed the asymptotic properties of the second stage estimator. The basic requirement is that the number of students N grows fast enough so that $\frac{J\ln(J)}{N}$ goes to zero, where J is the total number of schools (public and outside option) in the district. This ratio is approximately 0.069 in our application.

[28] This insight also suggests that lagged endogenous variables may be useful instruments and we explore these ideas as part of our sensitivity analysis. Dynamic panel data models often use timing assumptions and lagged endogenous variables to generate instrumental variables (Arellano and Bond (1991)). We implement this estimator below as a robustness check.

[29] For one example, Epple, Romano, and Sieg (2012) discussed the magnitude of relocation within metropolitan areas and the magnitude of the costs associated with such relocation.

by the district's default school assignment rules are not likely to be correlated with the actual unobserved characteristics of public schools under these four assumptions.

6. Estimation results for parental demand for schools

We estimate our empirical model of parental demand for schools with and without school-year fixed effects. We also implement our fixed effect, panel data logit demand model with and without random coefficients. We estimate our panel data random coefficients demand model on the subsample of students who were forced to move in order to estimate both permanent unobserved tastes for different peer characteristics and moving costs based on students' previous school choices (the "initial conditions" concern discussed in Section 5). This "no initial conditions" subsample includes all students in kindergarten, all students moving into the district, and all students moving from a K–5 school to sixth grade for the school years including and after the year they were forced to move.

There are 31,684 unique students in the full sample and 13,535 unique students in the "no initial conditions" subsample. Table 2 of the paper reports the first stage parameter estimates and estimated standard errors for different specifications of our demand model. We see from comparing the likelihood function values from Columns 1 and 2 of Table 2 that the fixed effects model fits the data much better than the model without fixed effects. We also find that adding random coefficients to our specification improves the fit of the model.[30]

We do not separately identify observed and unobserved school characteristics in the first stage; thus, the first-stage estimates presented in Table 2 do not rely on the validity of our instrumental variables.

We next implement our second stage decomposition of estimated school qualities into observed and unobserved components using the instrumental variables (IV) strategy discussed above.[31] Table 3 reports the point estimates and estimated standard errors for this second stage estimator. We consider three different model specifications. First, we report OLS estimates which do not account for the endogeneity of the peer characteristics in Column 1 of Table 3. Second, we show the results from a simple panel data estimator that uses lagged endogenous peer characteristics in Column 2. However, our preferred model specification relies on the peer characteristics predicted by the district's default school assignment rules; we denote predicted peer characteristics "Hoxby and Weingarth instruments" and report the results from this specification in Column 3.

Overall, we find that most parameters in both the first and second stages are precisely estimated and that the sign of our parameter estimates are plausible. We plot the implied distributions of the random coefficients determining parents' unobserved

[30]One cannot simply compare the likelihood value from Columns 2 and 3 of Table 2 as these two models are estimated on different student samples; we estimate the panel-data conditional logit model for the initial conditions subsample for our finding that "adding random coefficients to our specification improves the fit of the model."

[31]The first stage of our IV regressions are reported and discussed in the online appendix, available in a supplementary file on the journal website, http://qeconomics.org/supp/592/supplement.pdf.

TABLE 2. First stage estimates.

	No Fixed Effects No Random Coefficients Full Sample		Fixed Effects No Random Coefficients Full Sample		Fixed Effects Random Coefficients Initial Conditions Sample	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
School Achievement × Student Achievement	1.3433	0.0374	1.3715	0.039	2.0507	0.0847
School Achievement × Student Black	−0.1861	0.0647	−0.2908	0.073	−0.7057	0.1446
School Achievement × Student FRL	0.5255	0.0651	0.6083	0.072	0.2870	0.1402
School Proportion Black × Student Achievement	0.7657	0.0368	0.6855	0.0428	0.5725	0.0803
School Proportion Black × Student Black	0.7657	0.0368	3.0924	0.0850	5.914	0.1634
School Proportion Black × Student FRL	−0.6481	0.0660	−0.6243	0.0825	−0.3603	0.1381
School Proportion FRL × Student Achievement	0.4941	0.0391	0.3075	0.0486	0.5090	0.0958
School Proportion FRL × Student Black	−1.8534	0.0716	−0.8734	0.0839	−0.4846	0.1634
School Proportion FRL × Student FRL	2.4717	0.0777	3.9485	0.0860	5.755	0.1720
School Proportion FRL × Student Suspensions	0.1143	0.0064	0.1359	0.0095	0.1402	0.02618
School Suspensions × Student Achievement	0.0150	0.0065	0.0293	0.0101	0.0243	0.0240
School Suspensions × Student Black	−0.1282	0.0136	−0.0615	0.0203	−0.0353	0.0457
School Suspensions × Student FRL	0.0809	0.0144	0.0613	0.0204	0.0425	0.0477
School Suspensions × Student Suspensions	0.0154	0.0010	0.0320	0.0012	0.0643	0.0043

(*Continues*)

TABLE 2. *Continued.*

	No Fixed Effects No Random Coefficients Full Sample		Fixed Effects No Random Coefficients Full Sample		Fixed Effects Random Coefficients Initial Conditions Sample	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Moving Costs	−4.4206	0.0296	−4.3259	0.0241	−4.177	0.0518
Moving Costs × Student Achievement	−0.2953	0.0186	−0.2766	0.0136	−0.2185	0.0316
Moving Costs × Student Black	−0.0736	0.0327	−0.1650	0.0250	0.1502	0.0524
Moving Costs × Student FRL	0.1992	0.0333	0.1882	0.0263	0.1319	0.0569
Moving Costs × Student Suspensions	0.0421	0.0039	0.0500	0.0027	0.0500	0.0086
Travel Times	−0.3716	0.0041	−0.4223	0.0033	−0.3747	0.0049
Travel Times × Student Achievement	−0.0017	0.0029	0.0003	0.0017	−0.0319	0.0030
Travel Times × Student Black	0.1118	0.0046	0.1092	0.0034	0.0652	0.0052
Travel Times × Student FRL	−0.1226	0.0047	−0.1217	0.0035	−0.0685	0.0055
School Achievement × Standard Normal					3.082	0.0997
School Proportion FRL × Standard Normal					3.941	0.1207
School Proportion Black × Standard Normal					3.396	0.0968
School Suspensions × Standard Normal					0.4447	0.047
Likelihood	−160,902		−136,777		−74,763	

TABLE 3. Second stage estimates.

	No Random Coefficients			Random Coefficients	
	OLS	Lagged Regressors 2SLS	Hoxby and Weingarth 2SLS	Lagged Regressors 2SLS	Hoxby and Weingarth 2SLS
School FRL	−3.478	−3.063	−3.891	3.294	−3.877
	(0.457)	(0.507)	(0.459)	(1.129)	(0.570)
School Black	−1.713	−1.707	−1.446	−5.233	−4.045
	(0.163)	(0.168)	(0.216)	(0.542)	(0.273)
School Achievement	−0.441	0.061	0.150	10.746	4.102
	(0.279)	(0.292)	(0.347)	(3.796)	(1.563)
School Suspensions	−0.171	−0.145	0.015	−0.118	−0.229
	(0.031)	(0.032)	(0.056)	(0.031)	(0.052)

Note: Both specifications control for elementary school and middle school fixed effects. Robust standard errors included in parentheses.

tastes for peer characteristics in Figures 3–6. We find empirical evidence in favor of a "similarity effect": students want to sort into schools with other students that are similar to them. For example, though we see from Figure 3 that the vast majority of students in the sample value schools with high student achievement, the coefficient estimates from Table 2 indicate that higher achieving students have a higher valuation of average school-level achievement than lower achieving students. How a student values the school-level proportion of Black students and the school-level proportion of free or reduced cost lunch (FRL) students differs significantly by the student's own race and FRL status. In particular, both distributions are bimodal. In particular, students on FRL tend to prefer schools with a high school-level proportion of FRL students and vice versa. Similarly, students on FRL tend to prefer schools with a high school-level proportion of FRL students and vice versa. Our empirical findings therefore support the similarity hypothesis; students prefer environments that allow them to interact with other students that are similar in observed characteristics.

Figure 6 presents our estimated distribution of the random coefficient on the school number of suspensions. A high number of 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. Empirically, we see a wide dispersion in the distribution of parents' taste for school-level average number of student suspensions centered roughly at zero; thus, some parents positively value a higher average number of suspensions while others negatively value a higher average number of suspensions.

We also find from Table 2 that parents are reluctant to change schools once they have made their initial choices, as reflected by our estimates of the moving costs. These moving costs are identified by the lagged observed school choices made by students. Students also face significant travel costs; school choices that force students to commute for longer times are not popular.

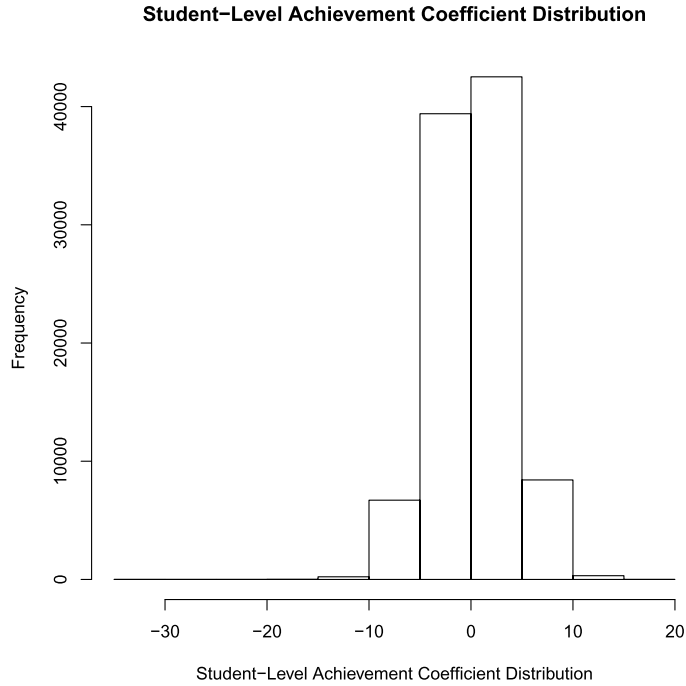

FIGURE 3. Estimated distribution of random coefficients: Achievement.

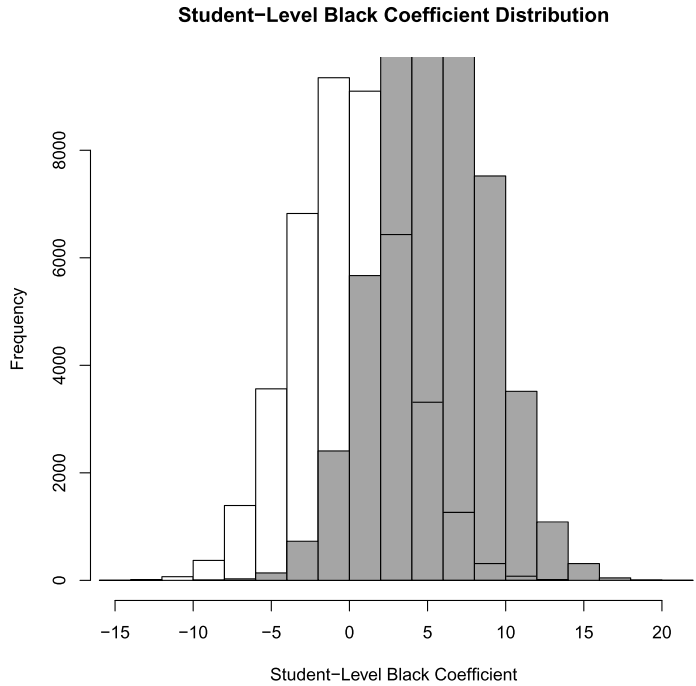

FIGURE 4. Estimated distribution of random coefficients: Race.

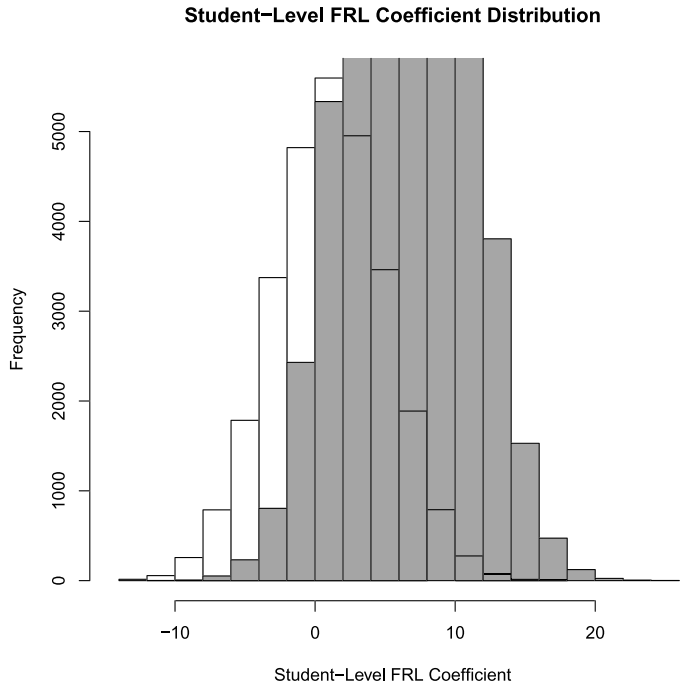

FIGURE 5. Estimated distribution of random coefficients: Poverty.

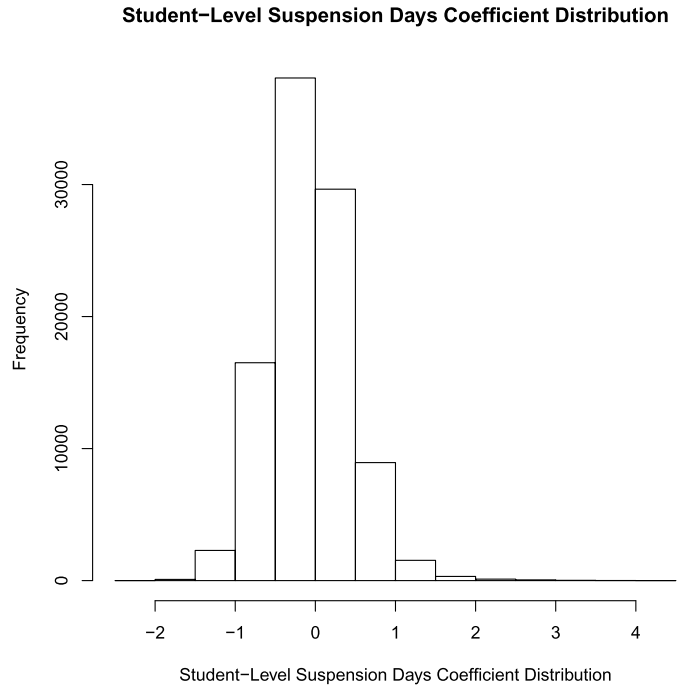

FIGURE 6. Estimated distribution of random coefficients: Suspensions.

TABLE 4. Estimated school quality and school closings.

	OLS	2SLS
Closed School	−1.693	0.150
	(0.344)	(0.178)
School FRL		−4.054
		(0.638)
School Black		−4.046
		(0.271)
School Achievement		3.746
		(1.682)
School Suspensions		−0.238
		(0.052)

Note: We use the panel-data random coefficients model for both specifications. We use the Hoxby and Weingarth instruments for the IV specification. Both specifications control for school year level fixed effects. Both specifications control for elementary school and middle school fixed effects. Robust standard errors included in parentheses.

We also estimate school-year fixed effects in our parental demand model for schools, which captures parents' perceived quality of schools in each school year. We test whether these school quality measures used by parents when making school choice decisions are similar to the quality measures used by the district. 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 between public schools that closed and public schools that remained open. Table 4 shows that our average estimated fixed effects for closed schools is significantly lower than the mean of the fixed effects of schools that remained open.

However, the first-stage school quality measures we estimate reflect both unobserved school characteristics such as teacher quality and observed peer characteristics across schools. We therefore perform a second hypothesis test for differences in school quality across public schools that closed versus remain open that conditions on observed peer measures and allows for potential time trends. We adopt a difference-in-difference style strategy in the second stage of the demand estimation procedure to implement this approach. Namely, we include the following variables to the second stage instrumental variable regression displayed in Table 4: (a) a post-reform time dummy; (b) an indicator for a school that was closed. Table 4 shows that parents do not perceive closed schools as systematically worse than the public schools that remained open once we condition on observed peer characteristics; parents perceive both sets of public schools as having similar unobserved school qualities. This reinforces our argument that accounting for student sorting is vital in determining which schools to close.

Finally, one may be concerned that our finding that students sort into schools with other students similar to them (i.e., a preference for homophily) may partially capture preferences over commuting distance. In particular, we may be concerned that parents choose to live in neighborhoods with children similar to their own and that most parents

TABLE 5. Distance and choices.

Variable	Prop. Attended School Within Quantile
Driving Times $\leq$ 25% Quantile of Driving Times	0.694
Driving Times $\leq$ 50% Quantile of Driving Times	0.795
Driving Times $\leq$ 75% Quantile of Driving Times	0.981

simply choose to send their children to the school closest to them. Table 5 reports that the fraction of students that attend schools within the 25th, 50th, and 75th percentile of driving times; we see from this table that almost 70% of all students attend a school that is less than or equal to 25% quantile of driving times.

We directly address this homophily versus preference for commuting time concern by reestimating our panel-data random coefficients demand model excluding all schools that are not within the 25% (50%) quantile of driving times from each student's choice set.[32] The Supplementary Material contains figures that plot the densities of parents' tastes for race and FRL for our original estimates and those obtained by using the 25% and 50% driving time definitions of the choice set. Overall, we find that the estimated distributions of tastes for both race and FRL do not vary substantially based on the definition of the choice set. The estimated coefficient that is most strongly affected by this definition of the choice set is the distribution of the travel time coefficient; this finding is not surprising since there is much less variation in travel times once we restrict the choice set to schools that are close in distance. Summarizing, our findings of homophily do not seem to be driven by preferences for commuting time; we find that students sort into schools with other students similar to them (i.e., a preference for homophily) even when choosing between schools with similar commuting times.

7. SCHOOL CLOSING ANALYSIS

The district closed a large number of underperforming schools after the 2005/06 school year; CCSD determined which schools to close by ranking them on a unidimensional quality measure based on students' standardized test scores. One key advantage of this approach is that it is straightforward to solve the school district's optimization problem even if a large number of schools need to be closed. The optimal policy is a cut-off rule; schools are closed in inverse order of their quality ranking until the required capacity reduction is achieved. We can compare outcomes implied by this closing rule with outcomes under alternative objectives in order to characterize the trade-offs faced by a superintendent.[33]

[32]We include the school that each student actually attended in the choice set (regardless of driving times) in order to obtain a well-defined likelihood function.

[33]As we noted previously, the superintendent very likely also considered subjective factors when making his school closing decisions; these subjective considerations may have resulted in deviations from closing schools strictly based on the performance ranking of schools. Our analysis of the "actual" school closing policy abstracts from the subjective judgments that are an inherent part of a superintendent's decision-making responsibilities.

The ideal policy experiment would be to compare the school closing agenda that was implemented by the district with the school closing scenarios predicted by our model under alternative objective functions. However, closing different sets of schools is a combinatorial problem; thus, considering all possible school configurations rapidly becomes computationally 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 for the number of schools actually closed by CCSD. We instead consider the case where the school district is tasked with closing down at least 5% of their overall public school capacity. We focus only on schools serving students in grades 6–8; this includes both middle schools (serving grades 6–8 exclusively) and K–8 schools. We restrict analysis to students in grades 6–8 in the school year 2005/06 immediately prior to the school closings; there are 8245 students in grades 6–8 in the district in 2005. We first consider the case where the district closes schools in inverse order of school quality ranking; for this objective function, we use the measure of school quality constructed by CCSD. We compare this "actual" scenario with closing down (at least) 5% of capacity under alternative objective functions.[34]

Table 6 summarizes the key findings of this paper. The table presents the school market outcomes from the optimal closing problem using the four different objective functions: (1) maximizing average school quality using the quality measures constructed by CCSD; (2) minimizing differences in peer characteristics across schools, that is, maximizing diversity; (3) minimizing the number of students not attending public schools in CCSD, that is, maximizing retention; (4) minimizing the number of students who relocate as a result of the school closings, that is, minimizing dislocation. As a shorthand, we refer to these four objective functions as quality, diversity, retention, and dislocation. We present results for two separate scenarios. First, we solve for the baseline probabilities implied by the sorting equilibrium where no schools are closed; we report magnitudes calculated using these baseline probabilities but not having students resort after schools are closed ("Pre-sorting School Market Outcomes"). We also report findings based on the probabilities calculated after schools are closed and students sort into the remaining schools ("Post-sorting School Market Outcomes"). Finally, all of the average and standard deviations reported in Table 6 are over the public schools that remain open after the school closings implied by each objective function.

Focusing on the post-sorting school market outcomes in the bottom panel of Table 6, we see from Row j that the school closings implied by the Retention objective in Column 4 result in the least number of students attending outside options (i.e., either private school or leaving CCSD); 1245.7[35] students attending outside options are even less than the 1444.3 students attending outside options if no schools are closed (Row j, Column 1). We see 1333.7 students attending outside options if we instead close down the schools

[34]Our school closing counterfactuals are based on our estimated demand model with school-year fixed effects, but without random coefficients in order to reduce computational burden (both the computational time associated with calculating the probabilities as well as adjusting these probabilities to account for school-level capacity constraints).

[35]We have noninteger numbers of students attending outside options as we are using conditional probabilities to assign students to schools.

TABLE 6. Optimal school closing analysis.

		1	2	3	4	5
		Baseline	Quality	Diversity	Retention	Dislocation
Pre-sorting School Market Outcomes						
a	Enrollment: Closed Schools	0	559.2	330.72	477.94	63.02
b	Mean FRL	0.70	0.68	0.70	0.73	0.69
c	Mean Black	0.60	0.58	0.60	0.64	0.58
d	Mean Achievement	−0.80	−0.82	−0.48	−0.70	−0.24
e	Mean Suspensions	5.01	4.86	2.87	5.33	4.61
f	Std. Dev. FRL	0.26	0.26	0.23	0.23	0.23
g	Std. Dev. Black	0.31	0.31	0.29	0.29	0.29
h	Std. Dev. Achievement	1.92	1.99	1.53	1.84	0.96
i	Std. Dev. Suspensions	10.93	11.31	3.08	11.27	11.46
Post-sorting School Market Outcomes						
j	Enrollment: Outside Options	1444.3	1605.3	1511.1	1245.7	1333.7
k	Mean FRL	0.70	0.69	0.70	0.71	0.68
l	Mean Black	0.60	0.59	0.60	0.63	0.58
m	Mean Achievement	−0.80	−0.83	−0.47	−0.70	−0.24
n	Mean Suspensions	5.01	4.87	2.85	5.27	4.59
o	Std. Dev. FRL	0.26	0.26	0.23	0.24	0.23
p	Std. Dev. Black	0.31	0.31	0.29	0.29	0.29
q	Std. Dev. Achievement	1.92	1.99	1.53	1.85	0.96
r	Std. Dev. Suspensions	10.93	10.66	3.16	11.24	11.44

Note: All means and standard deviations are for the remaining open public schools. We use the baseline probabilities for the pre-sorting panel. This table is for students in grades 6–8; $N = 8{,}245$.

with the least number of students (i.e., minimizing dislocation, Column 5); this number is still slightly smaller than the baseline number of students leaving for outside options (1444.3) without any schools being closed but is markedly larger than the number of students leaving for outside options if we explictly maximize retention (1245.7). This finding highlights the importance of student sorting; simply closing down the schools with the least number of students does not minimize the number of students switching schools because of student sorting. Finally, both the exogenous Quality objective (Column 2) constructed by CCSD based on student test scores as well as the Diversity objective (Column 3) intended to reduce disparities across schools in peer characteristics result in a large number of students leaving for outside options (1605.3 and 1511.1 students, respectively). Neither the quality metric constructed by CCSD nor the diversity objective focus on student retention; it is thus not surprising that more students leave for outside options under these objectives.

The remaining rows in the bottom panel of Table 6 show the average and standard deviation over all open public schools of our school-level peer characteristics. Most of these averages and standard deviations are similar across objectives, reflecting the fact that most students displaced due to their school being closed simply re-sort into another public school in CCSD. However, we see from the standard deviation of peer characteristics for the Diversity objective function in Column 3 that this objective primarily focused on reducing the standard deviation of the school average number of suspensions

(Row r). This makes sense, given that our empirical demand estimates from Section 6 indicate that: (1) parents on average exhibit stronger preferences for the other three peer characteristics (FRL, race, and achievement), and (2) some parents prefer schools with higher number of suspensions while others prefer schools with lower number of suspensions (see Figure 6). The main take-away from this is that superintendents interested in a diversity objective should focus on peer characteristics that do not feature prominently in parents' preferences and are horizontal rather than vertical (i.e., parents do not agree on whether the characteristic is "good" or "bad").

The top panel of Table 6 displays the school market outcomes using the baseline probabilities; we construct these baseline probabilities by solving for a sorting equilibrium without closing down any schools. However, the magnitudes in the top panel of Table 6 do *not* reflect the student sorting that occurs after schools are closed for each of the objectives displayed in Columns 2–5. We thus interpret these top panel results as what a naive policymaker might expect to happen if he or she closes schools without taking student sorting into consideration. We see from Row a of Table 6 that the Disruption objective closes down schools with the least overall number of students by (63.02); the other objectives close down on between 330–560 students, noting again that a lower number of students closed upon pre-sorting (Row a) do not necessarily mean that more students are retained in public schools post-sorting (Row j).

Both pre-sorting and post-sorting school market outcomes suggest that we see huge improvements in average school achievement from closing down schools based on the Dislocation objective (compare -0.8 to -0.24 in either Row d or m). This large improvement in average school achievement for the Dislocation objective is largely mechanical. Namely, this objective results in the closing of low-enrollment K–8 schools that happen to have low average achievement; these lower achieving students sort into schools with larger enrollments or leave for outside options. This mechanical effect also plays a role in why the average achievement for the Diversity and Retention objectives is higher than the Baseline. However, comparing the empirical findings across these objectives provides more insight. For example, we see that an objective explicitly based on retention results in lower average achievement relative to a Diversity-based objective (-0.70 versus -0.47 in Row m). This is because higher achieving students tend to be harder to retain, so a superintendent focused solely on retention will target retention of relatively lower achieving students. In contrast, a superintendent focused explicitly on keeping a "diverse" set of students will try to keep both low-achieving and high-achieving students in equal measure. Finally, the most striking result is that the objective explicitly based on maximizing *pre-sorting* average achievement results in the lowest *post-sorting* average achievement. If we close down schools with large numbers of low-achieving students, these students will simply sort into the remaining open public schools. This will both mechanically bring down the school-level averages of achievement for the remaining public schools but also induce high-achieving students at these public schools to leave for outside options. In short, a parent's perception of each school's quality depends crucially on peer characteristics; these peer characteristics are determined based on how students sort into schools. Thus, any measure of school quality used to close

down schools must take student sorting into account, as closing down schools will result in a new sorting equilibrium with different peer characteristics and thus different parental preferences for each school.

Our second stage student sorting game potentially has multiple equilibria. We address this multiplicity of equilibria concern by taking the optimal school closings implied by each of our objectives, and resolving for a new sorting equilibrium starting at different initial values for the peer characteristics associated with school-level proportions of FRL and Black students. We try two different sets of initial values: one where the proportions of FRL and Black students are both zero and one where the proportions of FRL and Black and students are both one. The resulting sorting equilibria found when using either set of initial values is quite similar to the one found in Table 6; the results using these starting values are reported in the Supplementary Materials. The qualitative conclusions of our optimal school closing analysis remain exactly the same for different initial values of the peer characteristics used to solve for a sorting equilibrium. Thus, it is unlikely that there exist different sorting equilibria that result in vastly different quantitative and/or qualitative results than Table 6. Put another way, the potential for multiplicity of sorting equilibria does not seem to be a concern for our empirical formulation of the optimal school closing problem.

In summary, we find that the choice of objective function has important consequences for the number of students who leave CCSD for outside options (either a private school in CCSD or leaving the district). Superintendents with objectives such as "maximizing quality" or "maximizing diversity" face a particularly difficult challenge retaining students in light of student sorting; indeed, if we try to maximize school quality without accounting for student sorting, our simulations indicate that a larger number of higher achieving students (the very students we are trying to retain) respond by leaving for outside options relative even to other objectives not focused on school quality. 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 framework highlights the trade-offs among these different objectives and demonstrates the importance of parental school choice decisions to school closing policies. Our results also show the potential of our framework in aiding the district-level management of school capacity decisions.

The state of the art in computational methods limits the scale of the school closing problem that we are currently able to consider. However, ongoing research in both operations research and numerical/computational methods are constantly expanding the size and scope of feasible mixed integer nonlinear programming problems such as ours. It is worth noting that our district confronted a large-scale school closing problem because of years of neglect in the face of declining enrollments and rising excess capacity. The same is true of some of the other urban districts now confronting large-scale closing problems. Realizing the issues surrounding these large reductions in public school capacity, districts may be more likely 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 computationally tractable using the framework that we have developed.

8. Conclusions

Many large urban districts face declining enrollment due to a combination of decreases in population as well as increased competition from charter schools, private schools, and suburban public schools. Hence, these urban districts are often forced to downsize. This paper formulates a sequential game of managing a school district's student capacity in order to address this important policy issue. This game has two stages: a superintendent is tasked with closing down a percentage of public school capacity in the first stage, facing the constraint that students will self-sort into the remaining schools in the second stage. Empirically, we first estimate the parameters of parents' demand for schools. We find empirical evidence from our demand model that students sort based on homophily: students want to attend schools with other students who are similar to them.

Thus, consideration of student sorting is vital to the assessment of any school closing policy. We solve the optimization problem facing a superintendent tasked with closing down public school capacity in light of this student sorting. We find that closing schools inevitably leads to additional retention concerns; moreover, higher achieving students are more likely to leave for outside options such as private schools or suburban public schools after school closings. 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 based explicitly on retaining these students. 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 to 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 must be retained and reassigned on the basis of seniority among the remaining schools. As a consequence, school quality in the remaining schools will change for two reasons: a change in 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 quality and principal quality across schools and time are captured in our estimation by school-year specific fixed effects. Solving the optimal school closing allowing for endogenous teacher reassignment would be 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. Summarizing, it is important to provide districts with more sophisticated tools for addressing the problems associated with declining enrollments that can be expected to stretch well into the future for many urban school districts. Such tools also have the potential to help districts with growing enrollments select among potential locations for opening new public schools.

Finally, we would like to point out that there are other potentially interesting applications of our framework. For example, General Motors discontinued the production and distribution of its Oldsmobile division in 2004; Oldsmobile had been producing cars for 107 years prior to 2004. The closing of the Oldsmobile division presaged a larger consolidation of GM brands as well as the discontinuation of certain models during the

company's 2009 bankruptcy reorganization. The bankruptcy reorganization included the sale or discontinuation of Hummer, Saturn, and Saab. In addition, GM discontinued the Pontiac brand in an effort to focus on their four major brands (Buick, Cadillac, Chevrolet, and GMC). Downsizing the product line of a company such as General Motors is, from a purely mathematical perspective, similar to the problem of downsizing a school district that has excess capacity. Namely, General Motors faces some demand for each of its products; its customers will resort into other GM products or buy a product from a different company after GM discontinues some of its models. Therefore, GM must choose which models to discontinue recognizing this customer sorting effect. We believe that the techniques developed in this paper will turn out to be fruitful in studying these types of applications as well.

References

Abdulkadiroglu, A., J. Angrist, S. Dynarski, T. Payne, and P. Pathak (2011), "Accountability and flexibility in public schools: Evidence from Boston's charters and pilots." *Quarterly Journal of Economics*, 126 (2), 699–748. [488]

Abdulkadiroglu, A., P. Pathak, and A. Roth (2005), "The New York City high school match." *American Economic Review P & P*, 95 (2), 364–367. [488]

Abdulkadiroglu, A., P. Pathak, and A. Roth (2009), "Strategyproofness versus efficiency in matching with indifferences: Redesigning the NYC high school match." *American Economic Review*, 99 (5), 1954–1978. [488]

Abdulkadiroglu, A. and T. Sonmez (2003), "School choice: A mechanism design approach." *American Economic Review*, 93 (3), 729–747. [488]

Ahmed-Ullah, N., J. Chase, and B. Secter (2013), "CPS approves largest school closure in Chicago's history." Chicago Tribune. [484]

Altonji, J. G., T. E. Elder, and C. R. Taber (2005), "Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools." *Journal of Political Economy*, 113 (1), 151–184. [488]

Angrist, J., E. Bettinger, E. Bloom, E. King, and M. Kremer (2002), "Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment." *American Economic Review*, 92, 1535–1558. [488]

Angrist, J., E. Bettinger, and M. Kremer (2006), "Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia." *American Economic Review*, 96, 847–862. [488]

Arellano, M. and S. Bond (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Review of Economic Studies*, 58 (2), 277–297. [503]

Bayer, P., F. Ferreira, and R. McMillan (2007), "A unified framework for measuring preferences for schools and neighborhoods." *Journal of Political Economy*, 115 (4), 588–638. [485]

Bayer, P., R. McMillan, and K. Reuben (2004), "The causes and consequences of residential segregation: An equilibrium analysis of neighborhood sorting." Working paper. [502]

Bayer, P. and C. Timmins (2005), "On the equilibrium properties of locational sorting models." *Journal of Urban Economics*, 57, 462–477. [497, 502]

Belotti, P., J. Lee, L. Liberti, F. Margot, and A. Waechter (2009), "Branching and bound tightening techniques for non-convex MINLP." *Optimization Methods and Software*, 4, 597–634. [498]

Berry, S. (1994), "Estimating discrete-choice models of product differentiation." *Rand Journal of Economics*, 25 (2), 242–262. [485, 503]

Berry, S., J. Levinsohn, and A. Pakes (1995), "Automobile prices in market equilibrium." *Econometrica*, 63 (4), 841–890. [485]

Berry, S., J. Levinsohn, and A. Pakes (2004), "Differentiated products demand systems from a combination of micro and macro data: The new vehicle market." *Journal of Political Economy*, 112, 68–104. [485]

Berry, S., O. B. Linton, and A. Pakes (2004), "Limit theorems for estimating the parameters of differentiated product demand systems." *The Review of Economic Studies*, 71 (3), 613–654. [503]

Black, S. (1999), "Do better schools matter? Parental valuation of elementary school education." *Quarterly Journal of Economics*, 114 (2), 577–600. [486, 503]

Bravo, D., S. Mukhopadhyay, and P. Todd (2010), "Effects of school reform on education and labor market performance: Evidence from Chile's universal voucher system." *Quantatative Economics*, 1, 47–59. [488]

Burer, S. and A. Letchford (2012), "Non-convex mixed-integer nonlinear programming: A survey." *Surveys in Operations Research and Management Science*, 17, 97–106. [498]

Card, D., D. Dooley, and A. Payne (2010), "School competition and efficiency with publicly funded catholic schools." *AEJ-Applied Economics*, 2 (4), 150–176. [488]

Chakrabarti, R. (2008), "Can increasing private school participation and monetary loss in a voucher program affect public school performance? Evidence from Milwaukee." *Journal of Public Economics*, 92, 1371–1393. [488]

Cullen, J., B. Jacob, and S. Levitt (2006), "The effect of school choice on student outcomes: Evidence from randomized lotteries." *Econometrica*, 74 (5), 1191–1230. [488]

Dinerstein, M. and T. Smith (2015), "Quantifying the supply response of private schools to public policies." Discussion Papers 15-019, Stanford Institute for Economic Policy Research. [488]

Dowdall, E. (2011), "Closing public schools in Philadelphia lessons from six urban districts." PEW Charitable Trust Philadelphia Research Initiative. [484]

Engberg, J., B. Gill, G. Zamora, and R. Zimmer (2012), "Closing schools in a shrinking district: Do student outcomes depend on which schools are closed?" *Journal of Urban Economics*, 72, 189–203. [484]

Epple, D., D. Figlio, and R. Romano (2004), "Competition between private and public schools: Testing stratification and pricing predictions." *Journal of Public Economics*, 88, 1215–1245.

Epple, D., R. Romano, and H. Sieg (2012), "The intergenerational conflict over the provision of public education." *Journal of Public Economics*, 96 (3), 255–268. [503]

Epple, D. and H. Sieg (1999), "Estimating equilibrium models of local jurisdictions." *Journal of Political Economy*, 107 (4), 645–681. [488]

Ferreyra, M. (2007), "Estimating the effects of private school vouchers in multi-district economies." *American Economic Review*, 97, 789–817. [488]

Figlio, D. and C. Rouse (2006), "Do accountability and voucher threats improve low-performing schools?" *Journal of Public Economics*, 90, 239–255. [488]

Geyer, J. and H. Sieg (2013), "Estimating an equilibrium model of public housing with rationing." *Quantitative Economics*, 4 (3), 483–513. [494]

Goldberg, P. (1995), "Product differentiation and oligopoly in international markets: The case of the U.S. automobile industry." *Econometrica*, 63 (4), 891–951. [500]

Hastings, J., T. Kane, and D. Staiger (2009), "Paternal preferences and school competition: Evidence from a public school choice program." Working paper. [488]

Hastings, J. and J. Weinstein (2008), "Information, school choice and academic achievement: Evidence from two experiments." *Quarterly Journal of Economics*, 123 (4), 1373–1414.

Hoxby, C., S. Murarka, and J. Kang (2009), "Charter schools in New York City: Who enrolls and how they affect their students' achievement." Working paper. [488]

Hoxby, C. and G. Weingarth (2006), "School reassignment and the structure of peer effects." Working paper. [485]

Hsieh, C. and M. Urquiola (2006), "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program." *Journal of Public Economics*, 90, 1477–1503. [488]

Hurdle, J. (2013), "Philadelphia officials vote to close 23 schools." New York Times. [484]

Kesten, O. (2010), "School choice with consent." *Quarterly Journal of Economics*, 125 (3), 1297–1348. [488]

McFadden, D. (1974), "Conditioanl logit analysis of qualitative choice behavior." In *Frontiers of Econometrics* (P. Zarembka, ed.), Academic Press, New York. [500]

McFadden, D. (1981), "Econometric models of probabilistic choice." In *Structural Analysis of Discrete Data With Econometric Applications*, MIT Press, Cambridge. [500]

Neal, D. (1997), "The effects of catholic secondary schooling on educational achievement." *Journal of Labor Economics*, 15 (1), 98–123. [488]

Neilson, C. (2013), "Targeted vouchers, competition among schools, and the academic performance of poor students." Working paper. [488, 492]

Rouse, C. (1998), "Private school vouchers and student achievement: An evaluation of the milwaukee parental choice program." *Quarterly Journal of Economics*, 113, 553–602. [488]

Rust, J. (1987), "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher." *Econometrica*, 55 (5), 999–1033. [498]

Skrainka, B. S. and K. L. Judd (2011), "High performance quadrature rules: How numerical integration affects a popular model of product differentiation." Available at SSRN 1870703. [501]

Su, C. (2014), "Estimating discrete-choice games of incomplete information: Simple static examples." *Quantitative Marketing and Economics*, 12 (2), 167–207. [487]

Su, C. and K. Judd (2012), "Constrained optimization approaches to estimation of structural models." *Econometrica*, 80 (5), 2213–2230. [487]

Witte, J., D. Carlson, J. C. D. Fleming, and P. Wolf (2012), *Voucher Longitudinal Educational Growth Study Fifth Year Report*. SCDP Milwaukee Evaluation, Report 29. [488]

Wooldridge, J. M. (2005), "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity." *Journal of applied econometrics*, 20 (1), 39–54. [501]

Co-editor Rosa L. Matzkin handled this manuscript.

Manuscript received 3 July, 2015; final version accepted 17 April, 2017; available online 22 June, 2017.