

Charter School Entry and School Choice: The Case of Washington, D.C.

**Maria Marta Ferreyra
Carnegie Mellon University**

**Grigory Kosenok
New Economic School**

October 25, 2011¹

*Preliminary and Incomplete
Please do not cite without permission*

Abstract

We develop and estimate a structural model of charter school entry. In this model, charter schools offer an alternative to neighborhood public schools and private schools. In the model's first stage, the potential entrant chooses his/her physical location and characteristics based on the perceived opportunities not exploited by conventional public and private schools. In the second stage, students choose among the available schools. We estimate the model using data for the Washington, DC school district. We use our parameter estimates to investigate the potential effects of changes in charter school legislation on charter school entry, and on post-entry outcomes such as enrollment and achievement.

¹ We thank Dan Akerberg, David Albouy, Tim Derdenger, Paul Ellickson, Dennis Epple, Liran Einav, Fernando Ferreira, Steve Glazerman, Brett Gordon, Lutz Hendricks, Susanna Loeb, Dan McMillen, Alvin Murphy, Aviv Nevo, Parag Pathak, Holger Sieg, Alan Sorensen, Chris Taber and Matt Turner for useful conversations and comments. We benefitted from comments by session participants at the 2009 LACEA conference, the 2010 Regional Science Association meetings, the 2011 AEFA conference, the 2011 SED conference and the 2011 CURE conference. Jeff Noel and Naomi Rubin DeVaux from FOCUS, Marni Allen from 21st Century School Fund, Steve Glazerman and the charter principals we interviewed for this study answered many of our questions on charters in DC. We thank our research assistants Gary Livacari, Sivie Naimer, Hon Ming Quek, and especially Nick DeAngelis for their help with the data collection. Jeff Reminga assisted us with the computational aspects of the project, and Bill Buckingham from the Applied Population Lab at the University of Wisconsin provided Arc GIS assistance. All errors are ours.

1. Introduction

The dismal academic performance of many students in central city school districts has been a source of growing concern in recent decades as academic achievement and high school completion have become increasingly important for economic success. Greater school choice has been advocated by many economists as a mechanism to foster improved achievement, with educational vouchers being the most prominent such proposal. While educational voucher programs have made significant inroads in some states and districts, voucher initiatives have faced intense opposition from teachers' organizations and have had limited success at the ballot box. By contrast, charter schools have received surprisingly widespread acceptance. This paper studies charter school entry and households' school choice, and focuses on Washington, D.C.

Charter schools are non-sectarian public schools of choice that operate free from many of the regulations facing ordinary public schools. They are publicly funded, cannot select their students, and cannot charge tuition. Unlike ordinary public schools (henceforth, public schools), charter schools are independent of the public school districts where they are located. To be admitted into a public school, a student usually has to reside within the geographic boundaries of the school's district and/or attendance zone. Charter schools, in contrast, do not have such requirements. If a charter school is oversubscribed, admission is determined by lottery.

Although charter schools do not enjoy the latitude envisioned by many voucher advocates, charter schools have considerable freedom in designing curricula and managing operations. For instance, they control the recruiting, retention and compensation of their teachers. Charters are regulated by state laws, and authorized by a chartering entity that approves the school's charter - the contract establishing the school's mission, program, goals, students served, and teaching philosophy. Depending on the state, the chartering entity is the school district, the state or some specific agency. In exchange for greater autonomy, charter schools are subject to the same accountability requirements as public schools, in addition to being overseen by the corresponding chartering entity. Although charter schools receive a per-student stipend, they rarely receive funding for facilities. Hence, the financial survival of charters is intimately connected with their ability to attract and retain students. While some charter schools have thrived and replicated their model across the nation, others have failed.

The charter school concept emerged in the late 1980's. It was embraced at its inception by Albert Shanker, then president of the American Federation of Teachers. Beginning with Minnesota in 1991, 40 states and the District of Columbia have passed charter laws. The institutional and regulatory environment facing charter schools varies widely across states. During the school year 2010-2011, there were about 5,400 charter schools in the country serving over 1.7 million students, or approximately 3% of the nation's primary and secondary education market.² While this aggregate market share for charter schools may seem small, large differences exist across states and cities in the nation, as explained below.

In order to open a charter school, a prospective charter operator must formulate and present a proposal to the chartering agency. This proposal sets forth the grade level(s) to be served and the thematic focus of the school. The latter includes a statement of subject emphasis (e.g., arts, language, math and science, core curriculum) and whether a special population is to be served (e.g., bilingual, disadvantaged, gifted). The proposal also includes a statement of whether a particular instructional model is to be implemented. For example, the applicant might propose to adopt an instructional model that is "franchised" nationwide, or a model that is being developed by the applicant. The proposal also includes anticipated enrollment, the physical facilities to be occupied, and a financial plan. In many ways, then, the charter applicant confronts a decision similar to that of a private firm contemplating entry into the market.

Most of the existing research on charter schools studies the effects of charter schools on the academic achievement of charter and public school students.³ This project, in contrast, seeks to study how charters come into existence in the first place, and how their presence shapes the marketplace for education. In particular, we address four questions. The first is what determines charter entry as well as the pattern of entry in terms of geographic location, curricular program offered, and grades served on the part of new entrants. These patterns are likely related to each other – for instance, in places with high-achieving public schools charters may be less likely to enter, and if entering they may be more likely to focus on a curricular specialization not offered

² See http://www.edreform.com/Fast_Facts/K12_Facts/.

³ See, for instance, Bettinger (2005) Bifulco and Ladd (2006), Booker et al (2007, 2008), Buddin and Zimmer (2005a, 2005b), Clark (2009), Hanushek et al (2007), Holmes et al (2003), Hoxby (2004), Hoxby and Rockoff (2004), Hoxby and Murarka (2009), Imberman (2009, forthcoming), Sass (2006), Weiher and Tedin (2002), and Zimmer and Buddin (2003).

by public schools. Our second research question is how students choose among public, charter and private schools. The set of options available to families differ depending on the age of their children, and their place of residence and income. It also varies over time as schools open or close. Even if all families faced the same options they might make different choices depending on their preferences, how they expect their child to perform in a given school, etc. Our third research question is how schools affect each others' decisions. For instance, the entry of charter schools may lead to the closing of some incumbent public and/or private schools – which, in turn, may further affect the remaining schools, including the new entrant. Our fourth research question is how charter entry, household choices and schools' interactions would be affected by changes in the regulatory and institutional landscape for public, private and charter schools. These counterfactuals are particularly timely given the current emphasis on charter schools in federal education policy.⁴ By modeling the interaction between charter entry, school competition, achievement and parental demand we can provide equilibrium, long-run policy answers. We can study actual programs and, perhaps more importantly, potential programs as well.

Addressing this interaction, however, is quite complex. To see why, consider the entry of a hypothetical charter. This provides families with a new choice, and some families may join the new school. In doing so, not only do they shape the composition of the student body of the new charter but they also affect the composition of the student body of their original schools. This alters the relative attractiveness of all schools, which further affects families' choices since parents care about schools' student bodies. A potential charter entrant must anticipate these effects in order to forecast the number of families that would ultimately attend the school.

This forecast is complicated by the fact that the entrant may not know exactly how good she is at starting a school from scratch. For instance, she must locate facilities (a particularly challenging activity in D.C.), hire a principal and teachers, appoint members to the charter school board, recruit students and parents, possibly request bank loans and make financial decisions, etc. Since these tasks may be considerably more involved than initially expected by the prospective entrant, she must forecast student demand for its services in the midst of uncertainty about her capacity to carry out the enterprise. Even after this uncertainty is resolved another

⁴ The federal "Race to the Top" program favors states with permissive charter legislation. See <http://www2.ed.gov/news/pressreleases/2009/06/06082009a.html> for further details.

source of uncertainty remains for this school – namely, how the actions of other schools will affect this particular school’s demand (similarly, the other schools are uncertain as to how this particular school will affect them).

Thus, we develop and estimate an equilibrium model of household school choice, charter school entry and school competition in a large urban school district. In our model, we view a charter entry point as a combination of location (neighborhood), thematic focus and grade level. Since charters must be able to forecast the demand for their services in order to assess their viability, we model how prospective entrants predict enrollment and demographic composition of their student body as a function of their geographic location, grades served, and thematic focus. The prospective entrant enters or not depending on the expected success of its entry and subsequent viability, which in our framework means maximizing expected net revenue. We model the entrant as being uncertain about its own quality at the entry stage.

We estimate the model using a unique and detailed data set from Washington D.C. from 2003 to 2007. The data set consists of information for all public, private and charter schools in the Washington, D.C. including enrollment by grade, school demographics, focus, and achievement. Our estimation exploits the variation across neighborhoods and over time in number, types and grade levels of schools, and enrollment and achievement. We treat each grade as a separate market, which has the benefit of increasing the effective sample size and generating exogenous variation in choice sets across households depending on the age of their children and their residential location. We cast the parental school model as an aggregate market share model and estimate it using an approach similar to Berry et al (1995), henceforth denoted BLP, and Nevo (2000, 2001). In the model, school characteristics such as the peer composition of the student body and aggregate achievement depend on households’ choices and their equilibrium sorting across schools. Our estimation strategy captures this aspect,⁵ which is critical to our ability to conduct counterfactuals.

⁵ General equilibrium analyses of school choice include Benabou (1996), Caucutt (2002), de Bartolome (1990), Epple and Romano (1998), Fernandez and Rogerson (1998), and Nechyba (1999, 2000). Calabrese et al (2006), Ferreyra (2007) and Ferreyra (2009) estimate general equilibrium models. Relative to these prior analyses, our econometric framework in this paper provides a model of choice among the entire set of schools within a district as well as a model of charter school entry. Our framing of peer effects is similar to the local spillovers model in Bayer and Timmins (2007), as explained below.

In modeling the charter school entry decision, we have chosen to focus on a single large urban district. This permits us to study the entry decisions of prospective entrants that all confront the same institutional structure, criteria for being granted a charter, reimbursement (funding) policies, and public school organization. Thus, we hold constant institutional structure in order to focus on the charter entrant and parental choice decision. We have chosen to focus on the Washington, D.C. school district, for the following reasons. First, the district has had a charter school law in place since 1996, and the law is highly permissive compared to charter school laws in other states (National Alliance for Public Charter Schools 2010). For instance, the per-student stipend for charter schools is equal to the full per-student spending in the Washington, D.C. school district, whereas it is equal to only a fraction of the corresponding district spending in most states. Moreover, in recent years a single entity, the Public Charter School Board (PCSB) has been responsible for the granting of charters. Second, in 2010 charter schools reached a student enrollment share of nearly 40 percent of all students in the public system (conventional public plus charter schools), making D.C. the district with the second largest charter market share in the nation (the first one being New Orleans).⁶ Third, the district is large, providing scope for potential entry of a large number of charter schools as well as substantial intra-district variation in population demographics across both residential locations and incumbent public schools. An analysis that includes multiple districts is challenging for a number of reasons—one being the challenge of obtaining a sufficiently complete data set—and we leave the pursuit of such a model to future research.

Charter school entry has been explored before by Glomm, Harris, and Lo (2005), who study charter school location in Michigan and California. They investigate the extent to which there is greater charter school entry where there is greater population heterogeneity. Using public school districts as the unit of observation, they find, as predicted, that there is more charter school entry where there is greater diversity in race and in adult education levels. In earlier work, Downes and Greenstein (1996) study private school entry across the 910 public school elementary districts in California. Rincke (2007) estimates a model of charter school diffusion in California. A theoretical model of charter school entry is developed by Cardon (2003), who

⁶ As of 2010, the districts where this share surpassed 30 percent were New Orleans, Louisiana (61 percent); Washington, D.C. (38 percent); Detroit, Michigan (36 percent); and Kansas City, Missouri (32 percent). Source: <http://www.charterschoolcenter.org>.

studies strategic quality choice of a charter entrant facing an incumbent public school. We build on the foundation established in these papers by modeling intra-district charter school entry decisions, parental choice, and the impact of entrants on public and private school incumbents. Perhaps closest to our approach is the work of Imberman (2009), who studies entry into a single large urban district in a reduced-form fashion, and Mehta (2011), who estimates a structural model of charter entry and competition between public and charter schools using data from North Carolina. The most salient differences between our work and Mehta's are the following: a) we endogenize school peer characteristics as equilibrium outcomes determined by household choices; b) we do not model the strategic interaction between charter and public schools;⁷ c) in our model, all charter schools in the economy are eligible to a given household regardless of its location, in accordance with the actual absence of residence requirements for charter schools.

Our work is also related to research on entry in empirical industrial organization. A recent review of this literature is provided in Draganska et al (2008). An important difference between our model and those found in the entry literature is that we do not model prospective entrants or incumbents as engaging in a game of strategic behavior. The assumption of non-strategic charter school entrants and incumbents is consistent with our understanding of the charter schools' decision processes. In particular, the primary competitors of charter school entrants in the district we study during our sample period are typically incumbent public and private schools rather than other charter schools.

Most entry studies rely on reduced-form models of demand for the calculation of profits. A recent exception is Draganska et al (2009), which incorporates a structural demand into a model of product choice and pricing. Our work differs in that our demand incorporates unobserved product characteristics as in BLP. We believe these to be particularly important in the case of schools, for which parents observe a number of aspects that are unobserved to the econometrician (e.g., some characteristics of the principal, teachers and facilities; culture of the school, etc.). In the model, some of these unobserved characteristics affect household utility whereas others affect student achievement (and, indirectly, household utility), although the lack

⁷ Between 1992 and 2007, the District of Columbia Public Schools (DCPS) had seven superintendents. This high turnover, coupled with financial difficulties and instability, suggest that DCPS did not react strategically to charters during our sample period.

of individual-level data leads us to treat them as an aggregate demand shock. Uncertainty over the permanent component of this shock is a key element facing the entrant at the time of submitting its application.

The recent entry literature has studied product choice and differentiation.⁸ Our study is related to this literature because it captures the variation in entry across locations, focuses and grade levels. Nonetheless, it is different in that one school-level characteristic valued by parents – namely, the composition of the student body – is the outcome of parental choices and the resulting student sorting across schools. Although the entrant cannot control parental choices, by deciding whether to enter or not in its entry point it affects parental choice sets. Thus, in order to make its entry decision, the entrant must forecast parents' responses should it enter, responses that will determine the student body composition of the new school and further affect parents' choices. These endogenous product characteristics are similar to Ferreyra (2007)'s endogenous peer quality in schools. They are in the spirit of the local spillovers in Bayer and Timmins (2007) and the endogenous neighborhood characteristics in Bayer et al (2007), both of which are the aggregate outcome of individual-level choices. We extend these papers by matching data on the composition of the student body at the school level, thus exploiting a natural set of overidentifying conditions that arises from the model.

While much of the entry literature relies on cross-sectional data, some recent studies (Dunne et al 2009, Suzuki 2010) have drawn on panel data such as ours. Our panel provides us with variation over time in entry patterns. Perhaps more importantly, by providing us with post-entry outcomes, the panel allows us to learn about the school-specific quality of both entrants and incumbents.

To estimate the parameters of household preferences and student achievement, an ideal dataset would include a panel of student-level observations with data on household residential location, school choice, and achievement. In the absence of such data, we rely on detailed demographics for the 433 block groups in Washington, D.C. For each block group we have estimated the joint distribution of child age, race, poverty status and family income. In addition, we observe the neighborhood public schools to which households in each block group are

⁸ See Draganska et al (2009) for product choice, and Mazzeo (2002), Seim (2006) and Suzuki (2010) for location choice, and the references therein.

assigned for each grade level. This provides us with a reasonable approximation to the actual distribution of choice sets available to households in different locations, and to households of different demographic profiles and child age within a given location. While BLP and Petrin (2002) also exploit aggregate Census data in order to estimate demand-side parameters, we rely on Census data aggregated at a much lower level.

We use our parameter estimates to study the effect of changes in the regulatory, institutional and demographic environment on charter entry, household sorting across schools and student achievement. For instance, we explore whether more permissive charter legislation would spur the creation of more charter schools, where these would locate, which students would attend them, and how achievement would change among the pre-existing schools.

The rest of the paper proceeds as follows. Section 2 describes our data sources and basic patterns in the data. Section 3 presents our theoretical model. Section 4 describes our estimation strategy. In Section 5 we provide some discussion and describe our intended counterfactuals. Section 6 concludes.

2. Data

In this section we describe the construction of our dataset and some salient aspects of the data, and direct readers interested in further detail to Appendix I. Our data set consists of information on every public, charter and private school in Washington, D.C. between 2003 and 2007. We have 1,243 campus-year observations. We have focused on the 2003-2007 time period to maximize the quality and comparability of the data over time and across schools. In addition, 2007 marked the beginning of some important changes in DCPS and hence constitutes a good endpoint for our window of study.⁹

2.1 Public and Charter Schools in Washington D.C.

⁹ In 2007, Michelle Rhee began her tenure as chancellor of DCPS. She implemented a number of reforms, such as closing and merging schools, offering special programs and changing grade configurations in some schools, etc.

In Washington, D.C., public schools fall under the supervision of the District of Columbia Public Schools (DCPS). Although there is only one school district in the city, there are many attendance zones. As for charters, until 2007 there were two authorizers: the Board of Education (BOE) and PCSB. Since 2007, the PCSB has been the only authorizing (and supervising) entity. An overarching institution at the “state” level is the Office of State Superintendent of Education (OSSE), which, among other roles, collects some data both for public and charter schools.

The unit of observation in our dataset is a campus-year (public and private schools have one campus each, but many charters have multiple campuses).¹⁰ Note that student-level data are not publicly available. We have 701 observations for public schools, and 230 for charters. For each school and year we observe address; enrollment by grade;¹¹ percent of white, black, hispanic and other-ethnicity students; percent of low-income students (i.e., students who qualify for free or reduced lunch); reading proficiency rate (i.e., the percent of students who are proficient in reading based on DC’s own standards and assessments), and math proficiency rate. Proficiency rates constitute our measures of achievement. For the few public schools that closed or merged during the sample period, we observe the year of closing or merger.¹²

In addition to the variables listed above, for charter schools we observe the authorizer, opening year and mission statement. For the charters that closed or merged in or after 2003, we observe the year of and reason for closing or merging.

The starting point for our dataset is the list of audited enrollments by school, year and grade from OSSE. This list includes public and charter schools. From this list we eliminated alternative and special education schools, early childhood centers, and schools that offered a residential program.¹³ Since OSSE also reports proficiency rates for each school and year as

¹⁰ A campus is identified by its name and not its geographic location. For instance, if a campus moves but retains its name, then it is still considered the same campus. In multi-campus charters, each campus has a separate address.

¹¹ We do not include adult or ungraded students, who account for less than 0.6% of the enrollment in this dataset.

¹² Washington, D.C. has a program of public school choice that allows students to enroll outside of their attendance zone. Since only 17 percent of students participated in this program during our sample period, we disregard it and leave its treatment for future extensions.

¹³ An alternative school is a school for children with behavioral problems. Children spend either the full year or part of it there. Sometimes they choose to attend these schools, but they are often sent there. These schools often enroll students aged 16-24. We excluded early childhood centers as long as they never had enrollment in regular grades during the sample period.

mandated by No Child Left Behind, our achievement data for public and charter schools come from OSSE. In 2003 and 2004, proficiency levels were determined according to the Stanford-9 assessment. To be considered proficient, a student was supposed to score at the national 40th percentile or higher. Since 2005, proficiency has been determined according to DC Comprehensive Assessment System. Before 2005, the grades tested were 3, 5, 8 and 10. Beginning in 2005, the grades tested have been 3 through 8 and 10. In compliance with NCLB, OSSE also reports the percent of White, Black, Hispanic and low-income students in the grades tested.

For public schools, the Common Core of Data (CCD) from the National Center for Education Statistics (NCES) is the main source for school address, ethnic composition and low-income status of the student body. When needed, we relied on OSSE's reporting of ethnic composition and low-income status. In addition, we drew on the Common Core of Data, web searches and phone calls to DCPS to identify year of closing or merger.

The fact that some charter schools have multiple campuses posed a major challenge in the construction of the dataset, as no single data source contains the complete or correct history of campuses. To re-construct this history, we drew on OSSE audited enrollments, School Performance Reports (SPR's) for PCSB-authorized charters, web searches of current websites and past Internet archives, charter school lists from Friends of Choice in Urban Schools (FOCUS), phone calls to charters that are still open, and achievement data at the campus level. The resulting campus-level data reflect our efforts to draw together campus-level information from these different sources, with the greatest weight being given to the OSSE audited enrollments, the SPRs and the NCLB achievement data. BOE charters that are no longer open represented a special challenge, as there is no institutional memory of these schools. Faced with conflicting information on campuses for these charters, in the absence of better evidence we took the most recent available addresses and campus configuration data as correct.

For PCSB-authorized charters, ethnic composition and low-income status come mostly from the School Performance Reports. For BOE-authorized charters, before 2007 this information comes from OSSE, and for 2007 from the School Performance Reports. If needed and possible, we supplemented these sources with the CCD.

We identified the year of opening based on SPR's, FOCUS listings and web searches. We identified the year of and reason for closing mostly based on reports from the Center for Education Reform and SPR's.

According to grades covered, we have classified schools into the following grade levels: elementary (if grades covered fall completely within the preschool-6th grade range), middle (if grades covered are 7th and/or 8th), high (if grades covered fall completely within the 9th - 12th grade range), elementary/middle (if grades span both the elementary and middle level), middle/high (if grades span both the middle and high level), and elementary/middle/high (if grades span the elementary, middle and high levels). This classification follows DCPS's criteria for identifying single-level categories (elementary, middle and high schools), and incorporates mixed-level categories, which are quite common among charters.

Each charter school has a statement that describes its mission or program focus. Mission statements come from the schools' web sites, SPR's, and FOCUS. Based on them we have classified thematic focuses into the following six categories: arts, civics, core curriculum (henceforth, core), language, math and science, and vocational.

2.2. Private Schools in Washington, D.C.

We have 312 school-year observations for private schools (private schools have one campus each). For each private school we observe address; enrollment by grade; percent of White, Black, Hispanic and other-ethnicity students; school type (Catholic, other religious, non-sectarian) and tuition by grade. Note that we do not observe proficiency rates or percent of low-income students for private schools.

With the exception of tuition, our data come from the Private School Survey (PSS) from NCES. The PSS is a biennial survey of private schools. We used the 2003, 2005 and 2007 waves. We imputed data for each school in 2004 by doing linear interpolation of the values of the corresponding variables between 2003 and 2005, and similarly for 2006. From the universe of schools surveyed by PSS, we eliminated alternative and special education schools, early childhood centers, schools with only ungraded students, and special academies.

We obtained tuition information for many private schools from their web sites. Though this information is current, in our empirical application we assume that relative tuitions among private schools have not changed since 2003.

2.3 Descriptive Statistics

The population in Washington, D.C. peaked in the 1950s at about 802,000, and declined steadily until 2000, when population was approximately equal to 572,000. The last ten years have witnessed some population gains, as the estimated 2009 population is approximately equal to 600,000. In particular, the population is estimated to have grown from 577,000 in 2003 to 586,000 in 2007, although the school-age population has actually declined from 82,000 to 76,000.¹⁴ The racial breakdown of the city has changed as well over the last two decades, going from 28, 65 and 5 percent White, Black and Hispanic in 1990 to 32, 55 and 8 percent respectively in 2007. Despite these changes, the city remains highly segregated by race and

¹⁴ Source: Population Division, U.S. Census Bureau. School-age population includes children between 5 and 17 years old. An alternative measure of the size of school-age population is total school enrollment in preschool through 12th grade, which has also declined from 87,000 to 81,000 students (see Figure 2).

income. In 2006, median household income was \$92,000 for whites, \$34,500 for blacks (Filardo et al, 2008).

2.3.1. Basic trends

In 2007, 58 percent of students attended public schools, 22 percent charter schools, and 20 percent private schools. In what follows, “total enrollment” refers to the aggregate over public, private and charter schools, and “total public” refers to enrollment in the public system (i.e., the aggregate of public and charter schools).

In national assessments, DC public schools have ranked consistently at the bottom of the nation in recent years. For instance, in 2009 the percent of 4th or 8th graders who are at or above the proficiency level in reading according to the National Assessment for Educational Progress was higher in all 50 states than in DC (similarly for math). Perhaps for this reason, charter schools have grown rapidly in DC since their inception in 1996. During our sample period alone, the number of charter school campuses almost doubled, going from 31 to 59, whereas the number of public and private school campuses declined slightly as a result of a few closings and mergers (see Figure 1 and Table 2). Over the sample period, 42 percent of private schools are Catholic, 25 percent are Other Religious, and 33 percent are Nonsectarian (see Table 6a).

Even though total enrollment has declined over time by about 6,000 students, enrollment in charter schools has risen approximately by the same amount (see Figure 2). As a result, the market share of charter schools has grown from 13 to 22 percent (see Figure 3), and charter share relative to total public enrollment has risen from 16 to 28 percent.

As Table 1a shows, private schools are demographically quite different from public and charter schools. For instance, 51 percent of private school students are white and 39 percent black, whereas about 5 and 83 percent of students in the public system are white and black, respectively. Moreover, private schools tend to be located in neighborhoods (Census tracts) with higher household income than public schools. This point is illustrated by Figures 4a through 4f, which depict the geographic location of public, private and charter schools at the beginning and end of our sample period.

A comparison of public and charter schools reveals that although Blacks and Hispanics account for most of the enrollment in the public system, charters tend to enroll proportionately more Blacks and fewer Hispanics than public schools. They also tend to enroll a higher fraction of low-income students. Moreover, they tend to locate in poorer neighborhoods. On average, achievement is quite similar among public and charter schools.

Public schools are highly heterogeneous, as are charters and private schools. Table 1b illustrates the heterogeneity of private schools. Catholic schools tend to locate in poorer neighborhoods than other religious or non-sectarian private schools. Whereas Blacks and Hispanics constitute almost 60% of Catholic school enrollment, 60% of the students in other private schools are white.

2.3.2. Variation by grade level

As column (1) of Table 3a shows, about two thirds of public schools are elementary. In contrast, 43 percent of charter campuses are elementary and 51 percent of private schools are elementary/middle. Furthermore, while only 4 percent of public schools mix grade levels, about a third of charter campuses and three quarters of private schools mix levels. In general, elementary schools are smaller in terms of enrollment than middle schools, which are in turn smaller than high schools. Nonetheless, at every level charters are smaller than public schools, and private schools are smaller than charters. The exception is high schools, as average enrollment in public and private schools is very similar yet higher than in charters. Different private schools, in turn, focus on different grade levels (see Table 3b). For instance, 77 percent of Catholic schools offer elementary/middle grades and only 15 percent offer high school grades, although the latter account for 42 percent of their total enrollment. In contrast, other religious and non-sectarian schools have very low enrolment in high schools, yet approximately 60 percent of their enrollment pertains to elementary/middle/high schools.

During the sample period, 55 percent of charter enrollment corresponds to preschool through 6th grade, 19 percent to grades 7 and 8, and 26 percent to grades 9 through 12. Importantly, market shares for public, private and charter schools differ greatly among individual grades (see Figures 5a, 5b and 5c), a feature of our data which we intend to exploit in the

estimation. For instance, private school shares grow steadily from prekindergarten up to 12th grade, thus peaking for high school grades at about 27 percent. Charter school shares peak for 7th grade, followed by 6th and 8th grade. Among the remaining grades, there is another charter share peak for kindergarten. During our sample period, charter market share has grown the most for early childhood and middle school grades. Charter shares peak for 7th grade because this is a natural entry point, since elementary schools in DCPS offer up to 6th grade and hence force children to switch schools for middle school. Moreover, many charters only take in new students in 7th grade. Note that the middle school peak for charters seems to absorb mostly students that would otherwise attend public schools.

It is quite common for charter schools to add grades over time until completing the grade coverage stated in the charter. For instance, a school's charter may state that in a few years the school will offer kindergarten through 5th grade, yet the school may open with kindergarten and first grade, add second grade the following year, third grade the year after, and so on. In fact, of the 30 campuses that entered between 2003 and 2006, 26 added grades during the sample period. While most additions were at the top (in the example above, adding second grade the second year of operations), some of them were at the bottom (as it would be the case if in the example above, the school added prekindergarten).

The variation in market share across grades by school type raises the question of whether the student body also varies across grade levels. The answer to this question is yes, as Table 4a shows. Much of this variation, however, reflects the variation in school type across levels. For instance, the high percent of white students in middle/high and elementary/middle/high schools reflects the fact that these (mixed) levels are mostly provided by private schools. Hence, Table 4b characterizes the student body at each level by school type. Among public schools, high schools have lower fractions of low-income students than other levels,¹⁵ and middle and high schools tend to be located in Census tracts with higher average income than elementary or elementary/middle schools.

¹⁵ This is consistent with the fact that the percentage of children who qualify for free- or reduced- lunch falls with age for D.C. as a whole. Nonetheless, it is interesting that the fraction of low-income children in charter schools is not lower for high school than for elementary or middle school.

Among public schools, reading and math proficiency decline with school level. In contrast, among charters reading and math proficiency peaks for middle schools. Interestingly, whereas public schools surpass charters in reading and math achievement at the elementary level, the reverse happens at the other levels. The gap between public and charter proficiency is greatest at the middle school level – while 36 percent of middle school students are proficient in math in public schools, 53 percent are proficient in math in charter schools.

Among private schools, almost two thirds of the enrollment in elementary/middle and middle schools is black, whereas about half of the enrollment in high schools is white. In elementary and middle schools Catholic schools have proportionally more black and Hispanic students than the other private schools, whereas the reverse happens for high school (see Table 6b). The higher the school level, the greater the prevalence of white students in all three types of private schools, which is consistent with the fact that the percent of white students decreases with school level in the public system.

2.3.3. Variation by charter focus

As column (1) of Table 5a shows, over the sample period 61 percent of the campuses offered a core curriculum. Among the remaining campuses, the most popular focuses are language, arts and vocational. However, this pattern of program specialization differs by level. While 56 and 74 percent of elementary and middle schools offer a core curriculum respectively, only 21 percent of high schools do. Among the mixed-level charters, elementary/middle and elementary/middle/high charters are almost exclusively devoted to a core curriculum, but only 21 percent of middle/high schools offer one. Moreover, language and arts are the more popular focuses after core among elementary schools, yet vocational and civics play this role among high schools.

As expected, different focuses attract different students (see Table 5b). For instance, language charters (usually focused on Spanish) tend to attract more Hispanic students. At the elementary and middle school level, math and science charters attract lower fractions of low-income students than the other focuses, though the reverse happens at the high school level. In fact, at the high school level, math and science have the highest fraction of low-income students.

At the middle and high school level, students in math-and-science charters outperform students in other charters, both in math and reading. If one excludes math and science, the highest average achievement corresponds to core schools at the elementary and middle-school level, and both to core and civics charters at the high school level.

While average achievement is quite similar for public and charter schools, the top-performing public and charter schools are quite different (see Table 9). The top-performing charter schools are located in much poorer Census tracts, enroll poorer students and are chosen by much larger fractions of African-American students than the top-performing public schools. Interestingly, they offer a core curriculum.

2.3.4. Charter relocations, closings and multiple-campus charters

A few campuses have relocated during our sample period (see Table 7). According to the SPR's, most of these campuses started out in a temporary location while they were renovating or preparing their permanent facilities, and then moved to their permanent facilities. The median distance move is 1.91 miles. Nonetheless, charters that plan to relocate usually inform parents of their plans ahead of time, so that the families can plan accordingly.

Furthermore, six campuses closed during our sample period, and six additional campuses from our sample have closed since 2008 (see Table 2). Of the closings that took place during our sample period, four were due to mismanagement and correspond to schools authorized by the Board of Education. The two other closings were due to academic reasons and correspond to schools authorized by PCSB.

As Table 8 shows, our sample includes 45 schools and 63 campuses. In general, multi-campus charters run multiple campuses for the sake of serving different grade levels. For instance, Friendship has two elementary school campuses (Southeast Academy and Chamberlain), one elementary/middle school campus (Woodridge), one middle school campus (Blow-Pierce), and one high school campus (College). The 10 multi-campus schools in our sample accounted for by 26 campuses and 46 percent of the charter enrollment during the sample period. Multi-campus schools tend to focus more on a core curriculum (70 percent of them do, v. 54 percent among single-campus schools) and to enroll slightly more African-Americans (89

percent of their students are African-American, relative to 82 percent among single-campus schools). A slightly higher fraction of their students is proficient in Reading or Math (39 and 37 percent in single-campus charters v. 40 and 40 percent respectively in multi-campus charters).

2.3.5. Early v. recent entrants

Given our focus on charter entry, an important question is whether the 27 campuses that entered before our sample period (“early entrants”) are systematically different from the 37 campuses that entered during our sample period (“recent entrants”). As Table 10 shows, these two groups are indeed different along a number of dimensions. Recent entrants tend to be smaller and more concentrated at the elementary or middle school level. They are also more likely to belong to a multi-campus charter organization. They enroll lower fractions of blacks and low-income students, and their achievement is lower on average. They are also less likely to offer a core curriculum.

To summarize, our data set is unique and draws from a variety of sources. It has not been compiled or used by any other researcher before. Moreover, it features a rich variation over time and across schools that will help us identify the parameters of our model. In the next section we formulate the model.

3. Model

In this section we develop our model of charter school entry and equilibrium. In this model, the economy is Washington, D.C. There are public, private and charter schools in the economy. Each school serves a different grade level, and there is a finite set of grade levels (elementary, middle, high, and combinations of these such as elementary/middle). The economy encompasses multiple neighborhoods or locations. At each point in time, each neighborhood has a prospective entrant for each grade level and focus. We use the term “entry point” to refer to each combination of location, grade level and focus. An incumbent charter school decides whether to remain open or exit, whereas a prospective charter school entrant chooses whether to enter or not. The prospective entrant forecasts demand and revenue in case of entering in its entry point,

and enters if the expected net revenue is positive.¹⁶ Public and private schools in the model are passive; the prospective entrant takes the locations, grades served, and focus of public and private schools as given (it also takes private school tuitions as given).

The economy is populated by households that live in different locations and have children who are eligible for different grades. For a given household, the school choice set consists of the public school assigned to the household by virtue of its residential location and the age of its child, and all private and charter schools in the economy that offer the child's grade, since private and charter schools have no residence requirements. For a given entry point, the prospective entrant forecasts its revenue based on a model of household choice that predicts the clientele that would be attracted in that entry point in competition with the incumbent public schools in that location, and all private and charter schools. The prospective entrant confronts uncertainty about its own quality in the form of an unknown shock that will affect demand if it should enter.

The model thus has two stages, an entry stage and a household choice stage. Since the latter is used in the former, we begin by presenting the model of household choice of school. Our school choice framework draws from Bayer and Timmins (2007).

3.1 School Choice and Market Equilibrium

The economy includes J schools, each one offering at least one grade. The economy is populated by households that have one child each. In what follows, we use “household”, “parent”, “child” and “student” interchangeably. Student i is described by $(g, D, \ell, I, \varepsilon)$, where:

- g is the grade of the student. Our data covers 15 grades: preschool, pre-kindergarten, kindergarten, and grades 1st through 12th.
- D is a vector describing student demographics. In our empirical application this vector has 5 rows, each one storing a 0 or 1 depending on whether the household is Black, White, Hispanic, or Asian (default race is Other Races), and low-income (this indicator equals 1 if the student qualifies for free- or reduced lunch, and 0 otherwise).

¹⁶ A potential extension is to have a charter school authorizer that imposes a minimum threshold for net revenue to internalize the externalities that a failing school imposes on its students.

- $\ell \in \{1, \dots, L\}$ is the location of the household in one of the L possible neighborhoods of the school district. In our empirical application, each location corresponds to a census block group. The location of a student determines her distance with respect to each school.¹⁷
- I is the income of the student's family.
- ε is a vector that describes the student's idiosyncratic preference for every school.

We use j and t subscripts to denote respectively a school¹⁸ and year. Our data includes $J=281$ schools and $T=5$ years (between 2003 and 2007). A household's choice depends on several variables that characterize a particular school and that are observed by the household at the time of making its choice:

- κ is the school's grade level, and thus denotes the set of grades served. A household chooses among the set of schools that offer the grade needed by its child.
- x_{ij} reflects exogenous characteristics of the match between the household and a school, such as the distance from the household to the school. In what follows we refer to x_{ij} as "distance."
- y_j denotes school characteristics chosen by the school, and whose choice we model for charter schools. One example is the school's thematic focus. Henceforth, for presentational clarity, we will refer to y_j as "focus." Denote by Y the number of possible focuses that charter schools can offer. We assume that public schools only offer one focus, a core curriculum. Private schools can offer one of three focuses (Catholic, Other Religious, nonsectarian). We do not model focus choice for public or private schools.
- p_{jgt} denotes school characteristics that we view as exogenous. One example is tuition in the case of private schools (tuition can vary by grade). Public and charter schools must charge zero tuition. In what follows, we refer to p_{jgt} as "tuition."

¹⁷ We assume that a student's location is given and does not depend on her choice of school. For models of joint residential and school choice, see Nechyba (1999, 2000) and Ferreyra (2007, 2009). In our empirical application, distance is measured as network distance, which can be expressed in miles or minutes (of travel time).

¹⁸ If the school has only one campus, j refers to the school; otherwise it refers to the campus. In this section we use "school" and "campus" interchangeably.

- \bar{D}_{jt} represents endogenous student peer characteristics of the student body at the school that are a result of aggregate household choices. In our empirical application this vector has 4 rows which store percent of Black, White, Hispanic students (default race is Other Races), and percent of low-income students respectively. These characteristics may change over time, as household choices change.
- ξ'_{jgt} is an unobserved (to us) characteristic of the school and grade. This includes characteristics of the teacher such as her responsiveness to parents and her enthusiasm in the classroom; physical characteristics of the classroom, etc.
- ξ''_{jgt} is an unobserved (to us) characteristic of the school and grade that affects children's achievement (in contrast, ξ'_{jgt} affects household satisfaction with the school and grade for reasons other than achievement). Thus, ξ''_{jgt} captures elements such as teacher effectiveness at raising achievement, the usefulness of the grade curricula to enhance learning, etc.

While y_j , p_{jgt} and \bar{D}_{jt} all represent school-specific characteristics, an important distinction is that y_j and p_{jgt} are chosen by the school, whereas \bar{D}_{jt} results from household choices.

Since we have enrollment data for each school, grade and year, we define a market as a grade. The size of the market for grade g in year t is M_{gt} , equal to the number of students who are supposed to enroll in grade g at time t .

The household indirect utility function is:

$$1. \quad U_{ijgt} = \delta'_{jgt} + \mu'_{ijgt} + \varepsilon_{ijgt}$$

where δ'_{jgt} is the common utility enjoyed by all the grade g children who enroll in school j at time t , μ'_{ijgt} is a student-specific deviation from the common school-grade utility, and ε_{ijgt} is an individual idiosyncratic preference for (j,g) at t . The common utility depends on school and peer characteristics as follows:

$$2. \quad \delta'_{jgt} = y_j \beta' + \bar{D}_{jt} \alpha' + \xi'_{jgt}$$

Here, α' and β' are vectors of parameters. In what follows, we refer to ξ'_{jgt} as a preference shock for school j and grade g at time t . At this point, a brief remark on notation may be helpful. Some components of the above utility function above are denoted with a prime, and some elements of achievement below are denoted with a double prime. This will permit us to economize on notation below when elements from the two components of the model are combined.

The household-specific component of utility is given by:

$$3. \quad \mu'_{ijgt} = E(A_{ijgt})\phi + D_i y_j \tilde{\beta}' + D_i \bar{D}_{jt} \tilde{\alpha}' + x_{ij} \gamma + \phi \log(I_i - p_{jgt})$$

This component of utility depends on the expected achievement of the student, $E(A_{ijgt})$, which is explained in greater detail below. It also depends on the interaction of y_j and D_i , which captures the variation in attractiveness of the thematic focus across students of different demographic groups, and the interaction of D_i and \bar{D}_{jt} , which captures the potential variation in preferences for school peer characteristics across different demographic groups. In addition, it depends on the distance between the household's residence and the school, and on school tuition.¹⁹

Student achievement depends on a school-grade factor common to all students, Q_{jgt} , a student's demographic characteristics, the fit of the thematic focus to the student (captured by the interaction of student demographics and focus), and a zero-mean idiosyncratic achievement shock v_{ijgt} , which parents do not observe at the time of choosing a school:

$$4. \quad A_{ijgt} = Q_{jgt} + D_i \omega'' + y_j D_i \tilde{\beta}'' + v_{ijgt}$$

As with most empirical analyses of schools, we do not have detailed measures of variables such as parental income, wealth or education. Differences in such variables across racial and ethnic groups motivate our inclusion of indicators for race and low-income status (eligibility for free or reduced lunch) in the achievement equation. As detailed below, the school-grade factor, Q_{jgt} , depends on the thematic focus of the school, peer characteristics of the student population, and a

¹⁹ As in BLP, our underlying utility function is Cobb-Douglas in disposable income after tuition, in characteristics of the school and in $\exp(\varepsilon_{ijgt})$. Taking logs of this underlying utility yields (1).

productivity shock ξ_{jgt}'' for school j and grade g at date t . Since peer characteristic measures are available at the school but not the grade level, below we do not place the subscript g on \bar{D} :

$$5. \quad Q_{jgt} = y_j \beta'' + \bar{D}_{jt} \alpha'' + \xi_{jgt}''$$

Substituting (5) into (4), we obtain

$$6. \quad A_{ijgt} = y_j \beta'' + \bar{D}_{jt} \alpha'' + D_i \omega'' + y_j D_i \tilde{\beta}'' + \xi_{jgt}'' + v_{ijgt}$$

Since parents observe ξ_{jgt}'' at the time they choose a school, their expectation of (6) is:

$$7. \quad E[A_{ijgt}] = y_j \beta'' + \bar{D}_{jt} \alpha'' + D_i \omega'' + y_j D_i \tilde{\beta}'' + \xi_{jgt}''$$

Substituting (7) into (3), we obtain:

$$8. \quad \mu_{ijgt}' = y_j \beta'' \phi + \bar{D}_{jt} \alpha'' \phi + D_i \omega + y_j D_i \tilde{\beta} + D_i \bar{D}_{jt} \tilde{\alpha} + x_{ij} \gamma + \phi \log(I_i - p_{jgt}) + \phi \xi_{jgt}''$$

where $\omega = \omega'' \phi$. The coefficient of the interaction of y_j and D_i is $\tilde{\beta} = \tilde{\beta}' + \phi \tilde{\beta}''$. This interaction captures both the variation in attractiveness of a school's focus across students of different demographic groups ($\tilde{\beta}'$) and the fit between focus and student type in the achievement function ($\phi \tilde{\beta}''$).

Substitute (2) and (8) into (1) and regroup terms to obtain:

$$9. \quad U_{ijgt} = \delta_{jgt} + \mu_{ijgt}' + \varepsilon_{ijgt}$$

where δ_{jgt} and μ_{ijgt}' are defined below in (10) and (12). We now turn to a discussion of these terms, beginning with the common utility component δ_{jgt} :

$$10. \quad \delta_{jgt} = y_j \beta + \bar{D}_{jt} \alpha + \xi_{jgt}$$

In this expression, the coefficient of y_j captures both household preference for school focus and impact of focus on achievement: $\beta = \beta' + \phi \beta''$. Thus, the model captures an interesting potential tension between school characteristics that enhance productivity and school characteristics that attract students. For example, a long school day may enhance achievement, but parents and students may not like the longer day. Similarly, the coefficient of \bar{D}_{jt} captures both household preference for peer characteristics and the impact of peer characteristics on student achievement: $\alpha = \alpha' + \phi \alpha''$. The error term in (10) impounds both a preference and a productivity shock:

$\xi_{jgt} = \xi'_{jgt} + \phi \xi''_{jgt}$. We will refer to this composite shock as a demand shock. Since the demand shock captures unobserved school-grade elements that affect both utility and achievement, it reflects the same kind of tension described above. For instance, parents may like the atmosphere created by a teacher in her classroom and the enthusiasm she instills in the students even if these are not reflected in higher achievement. Following Nevo (2000, 2001), we decompose the demand shock as follows:

$$11. \xi_{jgt} = \xi_j + \xi_g + \xi_t + \Delta \xi_{jgt}$$

In this decomposition, the school-specific component ξ_j captures elements that are common to all grades in the school and constant over time, such as the school's culture and characteristics of its facilities (provided facilities do not change over time). The grade-specific component ξ_g captures elements that are common to a given grade across schools and over time. For instance, grades have more classrooms in middle than elementary schools and are thus bigger. The time-specific component ξ_t captures shocks that are common to all schools and grades and vary over time, such as changes in accountability or testing requirements. Hence, $\xi_j + \xi_g + \xi_t$ is the mean unobserved characteristic for school j and grade g at time t , and $\Delta \xi_{jgt}$ is a deviation from this mean. This deviation is due, for instance, to changes of teachers over time for a particular grade and school. We assume that $E(\Delta \xi_{jgt})=0$.

The household-specific component of (9) is:

$$12. \mu_{ijgt} = D_i \omega + y_j D_i \tilde{\beta} + D_i \bar{D}_{jt} \tilde{\alpha} + x_{ij} \gamma + \phi \log(I_i - p_{jgt})$$

Since the household may choose not to send its child to any school, we introduce an outside good ($j=0$). The indirect utility from this outside option is:

$$13. U_{i0gt} = \phi \log(I_i) + \xi_{0gt} + D_i \omega_0 + \varepsilon_{i0gt}$$

Since we cannot identify ξ_{0gt} and ω_0 separately from the ξ_{jgt} terms of the “inside” goods or from ω , we apply the following normalizations: $\xi_{0gt} = 0$ and $\omega_0 = 0$.

Let J_{gt}^i denote the choice set of schools available to household i , who must send its child to grade g , at time t . Note that for a given grade, this choice set differs across households at a given time because of household residential location. It also differs over time for a given location

because of the entry and exit of schools that serve that grade. Let X_{ij} denote the observable variables that are either specific to the household or the match between the household and the school: D_i, I_i , and x_{ij} . A household that must send its child to grade g at time t chooses a school from the set J_{gt}^i in order to maximize its utility (it may also choose not to send its child to school). Assuming that the idiosyncratic error terms in (9) and (13) are i.i.d. type I extreme value, we can express the probability that household i chooses school j in grade g at date t as follows:

$$14. P_{jgt} \left(y_j, y_{-j}, \bar{D}_{jt}, \bar{D}_{-jt}, \xi_{jgt}, \xi_{-jgt}, p_{jgt}, p_{-jgt}, X_{ij}; \theta^d \right) = \frac{\exp(\delta_{jgt} + \mu_{ijgt})}{\exp(\varphi \log(I_i)) + \sum_{k=1}^{J_{gt}^i} \exp(\delta_{kgt} + \mu_{ikgt})}$$

where θ^d refers to the collection of demand-side parameters to be estimated. Let $h(D, \ell, g)$ be the joint distribution of students over demographics, locations and grades in the economy, and let $h(D, \ell | g)$ be the joint distribution of demographics and location conditional on a particular grade. Recall that each location ℓ is associated with a distance to each school. Given (14), the number of students choosing school j and grade g at time t is equal to:

$$15. N_{jgt} = \int \int_{\ell, D} P_{jgt}(\cdot) dh(D, \ell | g)$$

Thus, the market share attained by school j in grade g at time t is equal to:

$$16. s_{jgt} \left(y_j, y_{-j}, \bar{D}_{jt}, \bar{D}_{-jt}, \xi_{jgt}, \xi_{-jgt}, p_{jgt}, p_{-jgt}; \theta^d \right) = \frac{N_{jgt}}{M_{gt}}$$

Similarly, the equilibrium values of the student peer characteristics are as follows:²⁰

$$17. \bar{D}_{jt} \left(y_j, y_{-j}, \bar{D}_{jt}, \bar{D}_{-jt}, \xi_{jgt}, \xi_{-jgt}, p_{jgt}, p_{-jgt}; \theta^d \right) = \frac{\int \int_{\ell, g \in \kappa_j} \int_D D P_{jgt}(\cdot) dh(D, \ell, g)}{\int \int_{\ell, g \in \kappa_j} \int_D P_{jgt}(\cdot) dh(D, \ell, g)}$$

We will discuss how we match the empirical counterparts of the above quantities to those predicted by the model in Section 4. Note that in equilibrium, the school peer characteristics

²⁰ Although numerator and denominator add over all locations, note that in the case of public schools, a child who resides outside the school's attendance zone does not count the school in its choice set and hence cannot choose it. In other words, the child has zero probability of attending that particular public school.

taking as given by the families when making their school choices are the actual peer characteristics, as determined by families' choices.

Using the preceding and (6), achievement at the school level is:

$$18. A_{jt} = y_j \beta'' + \bar{D}_{jt} \rho + y_j \bar{D}_{jt} \tilde{\beta}'' + \bar{\xi}_j + \bar{\xi}_t + \bar{v}_{jt}$$

In this equation, $\rho = (\alpha'' + \omega'')$. The school fixed effect embodies the school's productivity shock and the mean of its grade-specific productivity shocks. The time fixed effect captures changes that affect achievement in all schools and grades, such as modifications to the assessment instrument. The error term captures school idiosyncratic productivity shocks and the mean of the idiosyncratic components of performance of the school's students. See Appendix II for a detailed derivation of (18).

3.2 *Charter Schools*

We distinguish between potential entrant and incumbent charter schools. Apart from the actual decision to enter, the main differentiating factor is that entrants are more uncertain about their demand shocks. Incumbent charters choose whether to remain in the market or exit, and whether to stay in the same location or move in case of remaining in the market. We do not model the entry or exit decisions of public or private schools.

3.2.1 *Entrants*

A few institutional details about charter school entry will provide a context for our modeling choices regarding charter entry. If a charter wishes to open in the Fall of year X, then it must submit its application by February of (X-1). The Washington, D.C. charter law specifies that the school's application must include a description of the school's focus and philosophy, targeted student population (if any), educational methods, intended location, recruiting methods for students, and an enrollment projection. The applicant must also file letters of support from the community, and specify two potential parents who will be on the school's board. In addition, the application must contain a plan for growth - what grades will be added, at what pace, etc.

At the time of submitting its application, the school must provide reasonable evidence that it will be able to secure a facility. The authorizer evaluates the enrollment projection by considering the enrollment in nearby public schools, similar incumbent charters, the size of the school's intended building, and how many students the school needs in order to be viable given the expected fixed costs.

If the application is approved, the charter receives approval notice in April or May of (X-1) and must start negotiations with the authorizer on a few issues, including the building. At the time of receiving the approval notice, the school should have secured a building, or else the negotiations with the authorizer will break down. Provided the school secures a building, it then uses the following twelve months to hire and train its prospective leaders, renovate the building (if needed), recruit students and teachers, and get ready to open its doors.

Charters are very aggressive in their efforts to recruit students. They do neighborhood searches, go to churches, contact parents directly, post flyers at public transportation stops and local shops, advertise in local newspapers and in schools that are being closed down or reconstituted, and host open houses. PCSB also conducts a "recruitment expo" in January, and charters participate in it. Word of mouth among parents also plays an important role. This is aided by the fact that the board of each charter must include two parents with children in the school.

Based on its projected enrollment, our hypothetical charter opening in Fall of X receives its first installment of funds in July of X. This means that any previous down payment on the facilities must be funded through a loan. An enrollment audit is conducted in October of X, and installments are adjusted accordingly.

Charters can run surpluses – this is the case, for instance, of charters that are planning on expanding in the future. They can also run deficits, as is the case with schools whose actual enrolment is too low relative to their fixed costs. However, PCSB only tolerates temporary deficits, and only in the case in which the school is meeting its academic targets. Thus, attracting and retaining students is of utmost importance to charters. Between 2004 and 2010, PCSB received 89 applications, of which 29 were approved.

Having laid down some institutional details of charter school entry, we turn now to the entry model. Recall that we have defined an entry point for charters as a combination of location

ℓ , focus y and grade level κ . We assume there is one prospective charter entrant in each entry point per period. The prospective entrant uses the preceding demand model to forecast its revenue in case of entering, and enters if the expected profit is non-negative. In the estimation of the supply side, we consider a neighborhood cluster as a prospective entry location.²¹

Recall that the demand shock for school j in grade g at time t can be decomposed as $\xi_{jgt} = \xi_j + \xi_g + \xi_t + \Delta\xi_{jgt}$. The school-specific component of this shock, ξ_j , is a permanent shock and reflects the “quality” of the school. Prospective charter entrants do not know their ξ_j before entering. In contrast, incumbents know their ξ_j , and these are public information. Both prospective entrants and incumbents observe ξ_g because it is time-invariant and common to all schools that offer g . Similarly, both prospective entrants and incumbents observe ξ_t because it is common to all schools in period t . At the beginning of t , neither entrants nor incumbents observe $\Delta\xi_{jgt}$.

In this one-period model, the timing of events in t is as follows:

1. At the beginning of t , each incumbent charter chooses whether to stay in its current location, move to another location, or exit from the market. Once they are made, incumbents’ decisions become public information.
2. The prospective entrant in each entry point chooses whether to enter or not. Recall that when making this decision, the prospective entrant does not observe its ξ_j or $\Delta\xi_{jgt}$.
3. At the end of t , households observe the demand shocks ξ_{jgt} of all schools, including new entrants and incumbents. Households choose schools based on this information.

We now provide further details on the information structure. As explained above, at the time of entry the potential entrant does not observe its school-specific “quality,” ξ_j , or its school-grade-time specific components $\Delta\xi_{jgt}$. With regards to ξ_j , all the entrant knows is that ξ_j is drawn from a normal distribution with mean μ_ξ and standard deviation σ_ξ . If the prospective entrant does indeed enter, it conducts a number of activities that enable it to learn ξ_j

²¹ A neighborhood cluster is a collection of Census tracts that is often used by DC planning agencies to proxy for a “neighborhood.” While D.C. includes 188 Census tracts, it only has 40 neighborhood clusters.

- it advertises its services, hosts open houses, hires a principal and teachers, participates in charter fairs, engages in fundraising, etc. These activities enable both the new charter and the parents to learn about the school’s “quality,” which is known to all parties thereafter. With regards to $\Delta\xi_{jgt}$, all the entrant knows is that it comes from a normal distribution with mean $\mu_{\Delta\xi}$ and standard deviation $\sigma_{\Delta\xi}$. We assume that ξ_j and $\Delta\xi_{jgt}$ are independent. Thus, at the beginning of period t the prospective entrant perceives ξ_{jgt} to be distributed normally with mean equal to $\mu_\xi + \xi_g + \xi_t + \mu_{\Delta\xi}$, and variance equal to $\sigma_\xi^2 + \sigma_{\Delta\xi}^2$.

When forecasting its market share, the prospective entrant observes ξ_j for all other schools and knows the decisions made by incumbents at the beginning of t . While the prospective entrant does not observe $\Delta\xi_{-jgt}$ for incumbents $-j$, it evaluates its forecast at $\Delta\xi_{-jgt} = 0$. We assume that the prospective entrant ignores the effect of other potential entrants when forecasting its own revenue.²²

Prospective entrant j in location ℓ , grade level κ and focus y computes its expected revenue in case of entering as follows:

$$19. E[s_{jgt} | (\ell, y, \kappa)] = \int_{\xi_{jgt}} s_{jgt}(y, y_{-j}, \bar{D}_{jt}, \bar{D}_{-jt}, \xi_{jgt}, \xi_{-jgt}, p_{-jgt}; \theta^d) N(\mu_\xi + \xi_g + \xi_t + \mu_{\Delta\xi}, \sigma_\xi^2 + \sigma_{\Delta\xi}^2)$$

where $-j$ indicates all incumbent schools, and ξ_{-jgt} is evaluated assuming $\Delta\xi_{-jgt} = 0$. Similarly, the prospective entrant forecasts its peer composition as follows:

$$20. E[\bar{D}_{jt} | y_j] = \int_{\xi_{jgt}} \bar{D}_{jt}(y_j, y_{-j}, \bar{D}_{jt}, \bar{D}_{-jt}, \xi_{jgt}, \xi_{-jgt}, p_{-jgt}; \theta^d) N(\mu_\xi + \xi_g + \xi_t + \mu_{\Delta\xi}, \sigma_\xi^2 + \sigma_{\Delta\xi}^2)$$

Let R_{gt} denote the reimbursement per student in grade g that a charter school obtains.²³ Let V_κ denote variable costs per student; these may differ by grade level, κ . Let ζ be an entry fee which

²² We adopt this assumption for two reasons. First, in this relatively early stage of the charter movement charters were mostly competing with incumbent public (and private) schools, as explained before. Second, entry is sufficiently rare in the data that ignoring other potential entrants may be a safe assumption.

²³ For a given year, this reimbursement varies across grades. For 2007, the base reimbursement (“foundation”) is equal to \$8002.06, and is adjusted by a grade-specific factor which is highest for high school and preschool. The

is only paid when entering. Let $F_{\ell t}$ denote fixed costs which must be paid every year that the school is open. These may vary by location and time, as the main component of fixed costs – the cost of facilities – varies across locations and over time.

The expected revenue net of costs of the prospective entrant j in location ℓ , grade level κ and focus y is

$$21. \pi_{\ell y \kappa t}^e(\theta^s) = \sum_{g \in \kappa} M_{gt} E[s_{jgt} | (\ell, y, \kappa)] (R_{gt} - V_{\kappa}) - \zeta - F_{\ell t} + v_{\ell y \kappa t}^e$$

where $\theta^s = \{\mu_{\xi}, \sigma_{\xi}, \mu_{\Delta\xi}, \sigma_{\Delta\xi}, \zeta, V_{\kappa}, F_{\ell t}, w\}$ refers to the collection of supply-side parameters to be estimated, and $v_{\ell y \kappa t}^e$ is measurement error in the charter school's profits that is unobserved by the econometrician. The prospective entrant enters if its expected revenue is non-negative. In the θ^s vector, parameter w denotes relocation costs and is used below in our model of incumbents. We assume that the utility of not entering is equal to v_{0t}^e . We assume that the error terms $v_{\ell y \kappa t}^e$ and v_{0t}^e are i.i.d. type I extreme value.

In practice, a charter school may serve only a single grade in the first year and grow into its full grade level over time. For example, a middle school that intends to serve 6th, 7th and 8th grade may enter serving only 6th grade in its first year, then adding 7th grade in its second year and 8th grade in its third year. Our characterization in (21) assumes that the school decides on entry based on whether it expects non-negative net revenues were it serving its full set of grades from the beginning of its operations.

3.2.2 Incumbents

Having previously entered, an incumbent school only faces uncertainty over the value of $\Delta\xi_{jgt}^{\xi}$. The incumbent knows that $\Delta\xi_{jgt}^{\xi}$ is distributed normally with mean $\mu_{\Delta\xi}$ and standard deviation $\sigma_{\Delta\xi}$. In addition, incumbent schools do not change their thematic focus or their grade levels. An

foundation is adjusted by inflation every year. In addition, charters in D.C. receive a facility allowance per child, equal to \$2,809.59 in 2007.

incumbent charter school decides whether it will continue operations, and if so, whether to stay in the current location or move to any of the other locations in the economy. Moving imposes cost w . When forecasting revenues in the current and alternative locations, the incumbent evaluates $\Delta \xi_{-jgt}^{\xi}$ at zero for other incumbents, and ignores the effect of potential entrants.²⁴ The incumbent also assumes that other incumbents will not move.

An incumbent charter school j located in ℓ , with focus y and grade level κ continues in operation at its current location, moves to a different location or exits entirely according to the outcome of the following optimization problem:

$$22. \pi_{jt}^i(\theta^s) = \max \begin{cases} \pi_{jt,stay}^i = \sum_{g \in \mathcal{K}} M_{gt} E[s_{jgt} | (\ell, y, \kappa)] (R_{gt} - V_{\kappa}) + F_{\ell t} + v_{jt\ell}^i & \text{if stays in location } \ell \\ \pi_{jt,move}^i = \sum_{g \in \mathcal{K}} M_{gt} E[s_{jgt} | (\ell', y, \kappa)] (R_{gt} - V_{\kappa}) + F_{\ell' t} - w + v_{jt\ell'}^i & \text{if moves to location } \ell' \\ \pi_{jt,exit}^i = v_{jt0}^i & \text{if exits} \end{cases}$$

where $v_{jt\ell}^i$ is measurement error in the incumbent's profits that is unobserved by the econometrician, and the expectations are taken over the distribution of $\Delta \xi_{jgt}^{\xi}$. We assume that $v_{jt\ell}^i$ and v_{jt0}^i follow an i.i.d. type I extreme value distribution.

4. Estimation

To estimate the model, we proceed in three stages. First, we estimate demand-side parameters θ^d . Second, we estimate supply-side parameters θ^s . Third, we estimate achievement-function parameters. Below we describe each stage.

4.1 Demand Estimation

²⁴ We make this assumption based on the fact that both incumbent moves and new entry are relatively infrequent in the sample.

We formulate household choice of school as a discrete choice problem and estimate preference parameters using an approach based on BLP. An important point of departure relative to BLP is our inclusion of endogenous peer characteristics of each school in household utility. BLP allows for endogeneity in prices but assumes all other observed product characteristics are exogenous. In our setting, school peer characteristics are the result of aggregate student choices, such as the fraction of a school's student body that is low income or a particular ethnicity. The peer characteristics are an equilibrium outcome yet simultaneously determine aggregate choices and represent an additional source of endogeneity. These endogenous school characteristics are similar to the local spillovers in the Bayer and Timmins (2007) model of sorting.

Since we view each grade as a market, our data consists of a panel of markets. To form the necessary moment conditions, we must first calculate a school's predicted enrollment share relative to market size. The paragraphs that follow describe this calculation and the computational implementation of the model.

For grade g and year t , market size M_{gt} is the number of children eligible for that grade. The main difficulty in calculating market size is that we do not have direct information on the number of children eligible for each grade in each year. Appendix III describes our solution to this problem and our calculation of market size and outside good share for each grade and year.

In the model, the economy is a collection of locations. For the sake of our demand estimation, a location ℓ consists of a Census block group (there are 433 block groups in D.C.), and each location is populated by households that are characterized by two features: the grade that its child must attend, and its demographic type. A household's demographic type m includes race (4 categories: White, Black, Hispanic, and Asian), income (16 values for income, each one representing the midpoint of the corresponding Census income bracket), and poverty status (2 categories: eligible for free- or reduced-lunch, and not eligible). Thus, we have $\bar{M} = 128$ ($=4*16*2$) household types for each of our 15 grades and 433 locations.

As for the household measures, using 2000 Census data we calculate the year-2000 number of households with children eligible for grade g in location ℓ that belong to demographic type m , or $\mu_{\ell gm}$ (see Appendix IV for further details on the construction of these measures). Note that $\mu_{\ell gm}$ varies across grades for a given location and demographic type; across demographic types for a given grade and location, and across locations, grades and demographic types as well.

Since we do not observe the joint distribution of location, grade, and demographic type for our sample period, in order to obtain the corresponding measures for years 2003 through 2007, we treat the year 2000 μ_{lgn} measures as a baseline and adjust them by a (grade, year) specific factor. This adjustment is meant to reflect the growth in the number of children eligible for each grade (a growth that potentially varies by grade and year). The adjustment relies on the assumption that while the marginal distribution of child age may change over time, the distribution of demographic type and location conditional on age (or grade) remains constant (see Appendix IV for additional details). In what follows we refer to μ_{lgmt} as year t 's measure of students for each location, grade, and demographic type.

As for school choice sets, we use attendance zone boundaries in D.C. to determine the public school to which each location ℓ is assigned for public schools at the elementary, middle and high school level (attendance zones are larger for middle and high schools than for elementary schools). These boundaries changed once during our sample period (in 2005), mostly to accommodate the public school closings and consolidations that took place in previous years.²⁵ Thus, for household of demographic type m sending its child to grade g and residing in location ℓ , at time t the choice set is given by the neighborhood public school offering grade g (identified through attendance zone boundaries as described above), and all private and charter schools in the economy offering grade g at time t .

Aggregating over all household demographic types for grade g , we can compute the predicted number of children enrolled in school j and grade g at time t as:²⁶

$$23. \hat{N}_{jgt} = \sum_{\ell=1}^L \sum_{m=1}^{\bar{M}} \hat{P}_{jgt} \left(y_j, y_{-j}, \bar{D}_{jt}, \bar{D}_{-jt}, \xi_{jgt}, \xi_{-jgt}, P_{jgt}, P_{-jgt}, X_{lmjt}; \theta^d \right) \mu_{lmgt}$$

where the set X_{lmjt} includes the variables that characterize households who live in location ℓ and belong to demographic type m in year t : the race, income and poverty status of household type m , and the distance from location ℓ to school j . The set X_{lmjt} varies over time because changes in the marginal distribution of grade (or age) affect the joint distribution of grade, location, and

²⁵ See Appendix IV for further details on the construction of choice sets.

²⁶ Note that for public schools, one does not need to aggregate over households in all location, given that only households that reside in the corresponding attendance zone are eligible to attend the schools.

demographic type. Denote by X_t the union of all the X_{lmj} sets for time t . Based on the above, the predicted enrollment share for (j,g) at t is equal to:

$$24. \hat{s}_{jgt} = \frac{\hat{N}_{jgt}}{M_{gt}}$$

Aggregating over all household types in school j we obtain the predicted equilibrium peer characteristics:

$$25. \hat{D}_{jt} = \frac{\sum_{g \in \mathcal{K}_j} \sum_{\ell=1}^L \sum_{m=1}^M D_m P_{jgt}(\cdot) \mu_{\ell gmt}}{\hat{N}_{jt}}$$

where D_m are the demographic characteristics of a household of demographic type m , and

$$\hat{N}_{jt} = \sum_{g \in \mathcal{K}_j} \hat{N}_{jgt}.$$

We do not expect the model to perfectly predict the school peer characteristics. Thus, the following relates the observed and predicted value of peer characteristics (\bar{D}_{jt} and \hat{D}_{jt} , respectively):

$$26. \hat{D}_{jt} = \bar{D}_{jt} + u_{jt}$$

where u_{jt} is estimation error, possibly due to sampling variation or measurement error. Let J^* be the number of school-year observations in our data, J^{**} be the number of school-grade-year observations, and D^* be the number of student body characteristics at the school level. Then, stacking the estimation error for all school-year observations and demographic characteristics yields the $(J^* \times D^*) \times 1$ vector u .

Denote by J^{**} the total number of school-grade-year enrollment shares, and by $\Delta\xi$ the corresponding vector of J^{**} school-grade-year specific deviations from mean unobserved utility. As in BLP, we assume the following moment condition:

$$27. E[\Delta\xi | Z^{**}, X] = 0$$

where X is the vertical stacking of all X_t matrices, and Z^{**} is a $J^{**} \times L$ matrix of instruments, with $L > \dim(\theta^d)$, that are assumed to be conditionally independent of the school-grade-time

specific deviations from the demand shock. The instruments in our application are critical because schools and households likely observe numerous factors that influence household school choice but are unobserved to us. We discuss our instruments in section 4.3. In addition, we have the following moment condition on the estimation error of peer characteristics:

$$28. E[u | Z^*, X] = 0$$

where Z^* is a $(J^* \times D^*) \times \bar{K}$ matrix of instruments. These conditional moments yield the following $2\bar{K}$ moment conditions:

$$29. E[Z^{**'} \Delta \xi] = 0$$

$$30. E[Z^* u] = 0$$

Denote by J_t the number of schools in the data in year t . Let κ_{jt} be the set of grades covered by school j in year t . The sample analogs of (29) and (30) are the following two $\bar{K} \times 1$ vectors:

31.

$$a) \quad q_s(\xi, Z^{**}) = \frac{1}{J^{**}} \sum_{t=1}^T \sum_{j=1}^{J_t} \sum_{g \in \kappa_{jt}} Z_{jgt}^{**'} \Delta \xi_{jgt}$$

$$b) \quad q_D(u, Z^*) = \frac{1}{J^*} \sum_{t=1}^T \sum_{j=1}^{J_t} Z_{jgt}^{*'} u_{jt}$$

where Z_{jgt}^{**} and Z_{jgt}^* are two $1 \times \bar{K}$ vectors of instruments.

To estimate the model, we reformulate the problem as a mathematical program with equilibrium constraints (MPEC). Dubé et al (2010), building on Su and Judd (2010), show how to estimate the aggregate demand model in BLP as a constrained optimization problem that simultaneously calculates the mean utilities and estimates the utility preference parameters. Thus, a key benefit in following the MPEC approach is that we avoid numerically inverting the market shares to obtain the mean utilities within a nested optimization, which significantly improves the speed and robustness of estimation. We briefly describe the method below and direct the reader to Dubé et al (2010) for details.

Let s be the vector that contains that J^{**} observed market shares, and let \bar{D} be the vector that contains the $J^* \times D^*$ observed peer characteristics. Let \hat{s} and $\hat{\bar{D}}$ be their predicted counterparts. Let q be the $2\bar{K}$ vector that stacks the q_s and q_D vectors in (31). The MPEC objective function minimizes an analog of a GMM objective function subject to a set of constraints:

$$\begin{aligned}
 & \min_{\theta^d, \xi, \lambda, u} \quad \lambda' W \lambda \\
 & \text{s.t.} \quad s = \hat{s}(\Delta \xi, \bar{D}; \theta^d) \\
 32. \quad & \bar{D} = \hat{\bar{D}}(\Delta \xi; \theta^d) - u \\
 & \lambda = q(\Delta \xi, u)
 \end{aligned}$$

where W is a weighing matrix. In contrast to the nested fixed point optimization in BLP, the MPEC algorithm simultaneously searches over values for the demand shocks $\Delta \xi$ and preference parameters θ^d . Given values for these, we can calculate the predicted market shares. The first constraint of the MPEC problem ensures the observed enrollment shares s match the enrollment shares predicted by the model given values for the preference parameters, demand shocks, and (observed) peer characteristics. The second constraint enforces the necessary equilibrium condition that observed peer characteristics match expected peer characteristics. Note that to calculate the predicted enrollment shares, we evaluate households' utilities at the observed peer characteristics (namely, in the first constraint we evaluate \hat{s} at \bar{D} rather than $\hat{\bar{D}}$). Thus, the second condition searches for parameter values that match the predicted peer characteristics with their observed counterparts. The third constraint is used to simplify computations by converting the objective function into a quadratic form in λ .

The decomposition of the demand shock in (11) suggests the inclusion of school-, grade- and time-fixed effects in the utility function. However, a school-specific dummy variable captures both the value of school characteristics that do not vary over time, $y_j \beta$, and the school-specific mean of unobserved quality, ξ_j . To estimate β and ξ_j separately, we apply a minimum-distance procedure as in Nevo (2000, 2001). Denote by B the $J \times 1$ vector of school-specific dummy variables; by y the $J \times Y$ vector of school focuses, and by ξ the $J \times 1$ vector of school-

specific demand shock. From (10) and (11), $B = y\beta + \xi$. Assuming that $E(\xi | y) = 0$, the estimates of β and ξ are $\hat{\beta} = (y'V_B^{-1}y)^{-1}y'V_B^{-1}\hat{B}$ and $\hat{\xi} = \hat{B} - y\hat{\beta}$ respectively, where \hat{B} is the vector of school dummy coefficients estimated from GMM, and V_B is the variance-covariance matrix of these estimates.

4.2 Supply Estimation

We estimate the supply side parameters θ^s through Maximum Likelihood Estimation. Denote by \tilde{I}_t the information available to the econometrician for events previous to t . This includes entry and exit episodes, school relocations, and characteristics of the incumbent schools as of the beginning of the sample period. The likelihood function is then

$$33. \tilde{L}(\theta^s) = \prod_{t=1}^T \left[\left\{ \prod_{j=1}^{J_t} \Pr(d_{jt} = \hat{d}_{jt} | \tilde{I}_t) \right\} \left\{ \prod_{\ell=1}^L \prod_{y=1}^Y \prod_{\kappa} \Pr(d_{\ell y \kappa} = \hat{d}_{\ell y \kappa} | \{d_{jt} = \hat{d}_{jt}\}_{j=1}^{J_t}, \tilde{I}_t) \right\} \right]$$

where T is the number of periods in the data, J_t is the number of incumbents at the beginning of period t , L is the number of locations, and Y is the total number of focuses that charters can choose from (recall that for the sake of supply-side estimation, a location ℓ is a neighborhood cluster).

As for the probabilities inside the likelihood, $\Pr(d_{jt} = \hat{d}_{jt} | \tilde{I}_t)$ is the conditional probability that incumbent j chooses the action \hat{d}_{jt} that is observed in the data. Recall that the number of possible actions for the incumbent is equal to $L+1$, as the incumbent can function in any of the L locations or exit. We assign $\ell = 0$ to the option of exiting. Thus, the probability that the incumbent chooses location $\hat{d}_{jt}, \hat{d}_{jt} \in \{0, 1, 2, \dots, L\}$ is given by

$$34. \Pr(d_{jt} = \hat{d}_{jt} | I_t) = \frac{\exp(\bar{\pi}_{j\hat{d}_{jt}}^i(I_t))}{\sum_{\ell=0}^L \exp(\bar{\pi}_{j\ell}^i(I_t))}$$

where $\bar{\pi}_{jt\ell}^i(I_t)$ is the mean profit of incumbent j at time t if it locates in ℓ . The profit is given by (22). The other probability, $\Pr\left(d_{\ell\kappa y} = \hat{d}_{\ell\kappa y} \mid \left\{d_{jt} = \hat{d}_{jt}\right\}_{j=1}^{J_t}, \tilde{I}_t\right)$, describes entry behavior in each entry point (ℓ, y, κ) . The probability of entering is given by

$$35. \Pr\left(d_{\ell y \kappa} = \text{enter} \mid \left\{d_{jt} = \hat{d}_{jt}\right\}_{j=1}^{J_t}, \tilde{I}_t\right) = \frac{\exp\left(\bar{\pi}_{\ell y \kappa}^e\right)}{1 + \exp\left(\bar{\pi}_{\ell y \kappa}^e\right)}$$

where $\bar{\pi}_{\ell y \kappa}^e$ is the mean profit from (21).

4.3 *Achievement Estimation, and Summary*

We estimate the parameters of the achievement function in (18) by regressing average achievement by school and year on school focus, demographics, the interaction between focus and demographics, and school and time fixed effects (Appendix II provides the derivation of (18)). Since the error term \bar{v}_{jt} includes an idiosyncratic productivity which varies by school, grade and time and is observed by the households when choosing schools, it is correlated with the school's demographic composition. Thus, we estimate (18) through instrumental variables.

Since not all the parameters of the model are identified (see the discussion below), there are no gains from estimating the achievement regression jointly with the other equations. We only use the parameters of the achievement function for our counterfactuals.

To summarize, our estimation approach is as follows. First, we match enrollment shares and student demographics in order to estimate the parameters of household preferences and unobserved school quality. Second, we match charter decisions in order to estimate the parameters of charter revenues and the distribution of unobserved quality of all charters. In this stage, we use the first-stage parameters to compute charter expected revenues. Third, we estimate the parameters of the achievement function.

4.4 *Identification and Instruments*

For the identification of the demand-side and achievement parameters, the main concern is the endogeneity of the student body composition in the utility and achievement functions. Much of this concern is alleviated by the inclusion of school-, grade- and time-specific dummy variables following the demand shock decomposition in (11). However, the concern remains that when households choose schools, they observe the school-grade-time specific deviation $\Delta\xi_{jgt}^{\xi}$, which we do not observe. This would induce correlation between student peer characteristics \bar{D}_{jt} , which are an outcome of household choices, and $\Delta\xi_{jgt}^{\xi}$.

We address this correlation by using the following instruments for \bar{D}_{jt} . For multi-campus charters, we instrument for the student body characteristics of one campus with the student body characteristics of the other campus/es in the same school. The idea behind this strategy is that campuses of the same school may be perceived as belonging to the same “brand,” yet the peer composition in one campus does not directly affect the utility of households in another campus, particularly in the case of campuses that serve different grade levels. This idea is similar to that of Nevo (2000, 2001), who instruments for the price of a brand in one market with the price of the same brand in other markets.

For single-campus charters, our instrument is the student body characteristics in similar but geographically distant charters, where “similarity” is defined in terms of focus and grade level.²⁷ For public schools, our instrument is the study body characteristics in public schools that serve the same grade level and are located in neighborhoods with similar demographics (for instance, in tracts with similar average household income) yet are geographically distant. For private schools, our instrument is the study body characteristics in similar yet distant private schools, where similarity is measured in terms of focus (Catholic, religious non-Catholic, nonsectarian), grades served, average tuition, and demographic characteristics of the neighborhood where they are located.

Assuming that these instruments are valid, we now turn to the identification of the achievement function parameters $(\rho, \beta^*, \tilde{\beta}^*)$ in equation (18). These are identified by the variation in school-level achievement, focus and demographics. Parameter β^* is the main effect

²⁷ This is also the strategy we use for multi-campus charters for the years in which they have only one campus open.

of school focus on student achievement while $\tilde{\beta}''$ is the effect of the match between student characteristics and the school focus (i.e., the coefficients of the interactions of school focus and student demographic characteristics). Recall that $\rho = \alpha'' + \omega''$. Thus ρ impounds the effects of students' own demographic characteristics, ω'' , and the effects of peer demographic characteristics, α'' , on achievement. Absent student-level data, these two effects of demographics cannot be distinguished. For the purpose of counterfactual policy analysis, however, knowledge of the combined effect, ρ , is sufficient. Having multiple observations per school identifies school-level fixed effects $\bar{\xi}_j$ in (18), and having multiple school observations per year identify the time fixed effects $\bar{\xi}_t$.

The parameters of the common component of utility, (α, β) in equation (10), are identified. Parameters α capture both the household preference for peer characteristics and the impact of peer characteristics on student achievement: $\alpha = \alpha' + \phi\alpha''$. As discussed below, parameter ϕ is not identified. This parameter captures the weight on achievement in the household utility function. We noted above that α'' is not identified. Hence, while α is identified, its individual components are not. Parameters β capture both the household preference for school focus and the impact of focus on achievement: $\beta = \beta' + \phi\beta''$. Hence, while β is identified and β'' is identified (as explained above), β' is not identified because ϕ is not identified. Again, however, from the perspective of counterfactual analysis of the impact of policies on school choice, identification of the components of α and β is not required. Having multiple observations of enrollment share for each grade and school allows us to identify the school-, grade- and time-fixed effects in the demand shock (ξ_j , ξ_g and ξ_t , respectively).

The parameters $(\tilde{\alpha}, \tilde{\beta}, \gamma, \omega, \varphi)$ of the household-specific component of utility (equation 12) are identified. These measure respectively the utility from the interaction of own and peer demographics, the utility from the interaction of own demographics and school focus, the utility of exogenous factors affecting the school-specific match (e.g., distance to school), the utility of the achievement impacts of a student's own demographic characteristics, and the utility of the consumption of all other goods. Recall that the utility of the achievement impact of a student's

own demographics is $\omega = \omega' \phi$. While ω is identified, its components are not. As discussed above, the lack of individual-level achievement data prevents us from identifying the achievement impact of own demographics, ω' . Hence, the weight of achievement on utility ϕ is not identified either. Recall that the coefficient on the interaction of own demographics and school focus is $\tilde{\beta} = \tilde{\beta}' + \phi \tilde{\beta}''$. While $\tilde{\beta}$ is identified by enrollment patterns across different school focuses, and $\tilde{\beta}''$ is identified as the achievement effect of focus, $\tilde{\beta}'$ (and ϕ) are not identified.

On the supply side, the entry fee ζ is identified by the frequency of entry in the data. The fixed cost F_{it} is identified by the variation in entry patterns across locations and over time, holding entrant and demand characteristics constant. Variable costs V_k are identified by the pattern of entry by grade level holding other things constant. Moving costs w are identified by the frequency of moves. Finally, the parameters of the distribution of demand shocks, $\mu_\zeta, \sigma_\zeta, \mu_{\Delta\xi}, \sigma_{\Delta\xi}$, are identified by the frequency of entry, exit and moves in the data, and by the distribution of ξ_j for the actual entrants (which we estimate based on demand-side fixed effects). Note, however, that we cannot estimate these parameters solely off the distribution of the estimated ξ_j 's for the actual entrants because those parameters pertain to the distribution of unobserved school quality for *all* charters, including those that do not enter.

5. Discussion and Intended Counterfactuals

In this section we first discuss some aspects of the model that might affect our parameter estimates, and then enumerate the counterfactual scenarios we will evaluate using our estimates.

For all its richness, our data is limited in some regards. These limitations, most of which are due to the unavailability of the corresponding data, have forced us to ignore certain institutional features of charter schools in our model. First, we do not observe charter capacity or effective demand (i.e., the number of students who apply to the school). Thus, our model cannot capture a distinctive aspect of charter schools, namely that they must randomize access in case of being oversubscribed. Second, we do not observe the complete set of charter applications; we

only observe the applications that were approved and entered the market. During the first part of our sample period there were two charter school authorizers in D.C., the BOE and the PCSB (see Section 2). Some have suggested that the BOE tended to authorize lower-quality applicants (Buckley and Schneider 2007). This, in turn, might have led prospective entrants to “shop” authorizers at the beginning of the sample period, and could have been reflected in a change in the distribution of school-specific quality when PCSB became the sole authorizer. Third, in reality charters choose whether to be single- or multi-campus organizations, in a choice that probably resembles a store chain’s decision to open a new franchise. We plan to conduct some sensitivity analyses aimed at learning about the impact of these constraints.

We plan to use our parameter estimates to conduct some counterfactuals. First, we will study the response of charter entry and student sorting to changes in per-student funding for charter schools. Given that real estate is a prime concern for charters, we are particularly interested in examining the consequences of raising the facilities allowance for charters. On a related note, we will study the response of greater access to facilities (represented as a lower fixed cost in some locations). DCPS has made some vacant public school buildings available for charters (Filardo et al, 2008). As public school enrollment continues to decline, in part for reasons unrelated to charters, the supply of facilities for charters should increase. Moreover, in recent years charters have had increasing access to “incubator facilities” where they are housed for a few years until moving to their permanent location. We can capture the greater access to initial facilities through lower entry fees and/or lower fixed costs for certain locations.

While many states provide free transportation for children (even for those attending private or charter schools), D.C. does not provide any busing for public, private or charter school children. Thus, the provision of publicly-funded busing could alter household choices significantly. It could also alter the geographic pattern of charter entry and location.

The charter landscape is heavily influenced by the preferences of the authorizer. Hence, changes in these preferences are likely to affect charter entry and student sorting. For instance, some claim that the authorizer today is less interested in approving vocational charters than it was a few years ago. Thus, it is of interest to study whether students would be less likely to attend charters if they were not of the exact focus that they prefer.

DCPS has undergone important changes in recent years. These changes include school closings, consolidations, re-configuration of grades, and adoption of specialized curricula. Moreover, public-school choice (with out-of-boundary enrollment) has become more prevalent. We will study the effect of these changes on charter entry and student sorting.

Washington, D.C. is home to a publicly-funded voucher program for private schools. Since the recipients of these vouchers are demographically similar to the students attending charters (Filardo et al, 2008), an expansion of the current program is likely to affect charter schools. Our model allows us to study this issue. Finally, one of the main demographic changes affecting most urban school districts in the United States is the loss of school-aged children. Thus, we will explore the response of charters and household to exogenous demographic shocks that change the potential enrollment in the city as a whole, or that change the income distribution of the families with school-age children.

6. Conclusion

In this paper we have developed a model of charter school entry and household choice of school and have devised a estimation strategy for the model. We will estimate the model using a unique dataset for Washington, D.C., which incorporates information on all public, charter and private schools in D.C. between 2003 and 2007. Since we rely on an equilibrium framework, we are able to capture the fact that the student body composition of each school is actually the outcome of parental choices, and that parents, in turn, respond to this composition when making choices. We model charter entrants as being uncertain about their school-specific quality, and making their entry decisions based on their expected revenue given the opportunities available to households.

Understanding the decisions made by charters and households helps us predict their responses to policy changes. Through our counterfactuals we will analyze alternative policies facing charter schools. Today, charter schools not only provide children with additional school choices but also provide researchers with new evidence on school management methods, educational curricula, and a number of aspects in which charters can diverge from public schools by virtue of the freedoms that have been granted to them. Thus, in future research we hope to explore the innovation and competition induced by charters in the education market.

References

- Bayer, Patrick and Christopher Timmins (2007). "Estimating Equilibrium Models of Sorting Across Locations," *Economic Journal*, 117(March), 353-374.
- Bayer, Patrick, Fernando Ferreira, and Robert McMillan (2007). "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, vol. 115, No. 4, 588-638.
- Benabou, Roland (1996). "Heterogeneity, Stratification and Growth: Macroeconomic Effects of Community Structure," *American Economic Review*, 86, 584-609.
- Berry, Steven, James Levinsohn, and Ariel Pakes (1995). "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841-890.
- Bettinger, Eric P. (2005). "The Effect of Charter Schools on Charter Students and Public Schools," *Economics of Education Review*, 24, 113-147.
- Bifulco, Robert and Helen F. Ladd (2006). "The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina," *Education Finance and Policy*, 1 (1), 123-138.
- Booker, Kevin, Scott Gilpatric, Timothy Gronberg, and Dennis Jansen (2007). "The Impact of Charter School Attendance on Student Performance," *Journal of Public Economics*, 91 (5/6), 849-876.
- Booker, Kevin, Scott Gilpatric, Timothy Gronberg, and Dennis Jansen (2008). "The Effect of Charter Schools on Traditional Public School Students in Texas: Are Children Who Stay Behind Left Behind?," *Journal of Urban Economics*, 64 (1), 123-145.
- Buckle, Jack and Mark Schneider (2007). *Charter Schools: Hope or Hype?*, Princeton and Oxford: Princeton University Press.
- Buddin, Richard and Ron Zimmer (2005a). "Is Charter School Competition in California Improving the Performance of Traditional Public Schools?," RAND Working Paper.
- Buddin, Richard and Ron Zimmer (2005b). "Student Achievement in Charter Schools: A Complex Picture," *Journal of Policy Analysis and Management*, 24 (2), 351-371.
- Calabrese, Stephen, Dennis Epple, Thomas Romer, and Holger Sieg (2006). "Local Public Good Provision: Voting, Peer Effects, and Mobility," *Journal of Public Economics*, 90, 6-7 (August): 959-81.
- Cardon, James H. (2003). "Strategic Quality Choice and Charter Schools," *Journal of Public Economics*, 87 (3), 729-737.
- Caucutt, Elizabeth (2002). "Educational Policy when There Are Peer Group Effects—Size Matters," *International Economic Review*, 43, 195-222

Clark, Damon (2009). "The Performance and Competitive Effects of School Autonomy," *Journal of Political Economy*, 117 (4), 745-783.

de Bartolome, C. (1990). "Equilibrium and Inefficiency in a Community Model with Peer Group Effects," *Journal of Political Economy*, 98, 110–133.

Downes, Thomas A. and Shane M. Greenstein (2006). "Understanding the Supply Decisions of Nonprofits: Modelling the Location of Private Schools," *RAND Journal of Economics*, Vol. 27(2), Summer, pp. 365-390.

Draganska, Michaela, Michael Mazzeo, and Katja Seim (2009). "Beyond Plain Vanilla: Modeling Joint Product Assortment and Pricing Decisions," *Quantitative Marketing and Economics*, 7(2), 105-146.

Draganska, Michaela, Sanjog Misra, Victor Aguirregabiria, Pat Bajari, Liran Einav, Paul Ellickson, Dan Horsky, Sridhar Narayanan, Yesim Orhun, Peter Reiss, Katja Seim, Vishal Singh, Raphael Thomadsen, and Ting Zhu (2008). "Discrete choice models of firms' strategic decisions," *Marketing Letters*, 19, 399–416.

Dubé, Jean-Pierre, Jeremy Fox, and Che-Lin Su (2010). "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation," Working paper No. 09-07, University of Chicago Booth School of Business.

Dunne, Timothy, Shawn Klimek, Mark Roberts and Daniel Yi Xu (2009). "Entry, Exit and the Determinants of Market Structure," NBER Working Paper 15313.

Epple, Dennis and Romano, Richard (1998). "Competition Between Private and Public Schools, Vouchers, and Peer-Group Effects". *American Economic Review*, March, 88(1), pp. 33-62.

Fernandez, Raquel and Richard Rogerson (1998).: "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform," *American Economic Review*, 88, 813–833.

Ferreya, Maria M. (2007). "Estimating the Effects of Private School Vouchers in Multi-District Economies". *American Economic Review*, 97 (3): 789-817.

Ferreya, Maria M. (2009). "An empirical Framework for Large-Scale Policy Evaluation, with an Application to School Finance Reform in Michigan". *American Economic Journal: Economic Policy*, vol. 1 no. 1: 147–80.

Filardo, Mary, Marni Allen, Nancy Huvendick, Ping Sung, David Garrison, Margery Turner, Jennifer Comey, Barika Williams and Elizabeth Guernsey (2008). "Quality Schools and Healthy Neighborhoods: A Research Report." Available at www.brookings.edu.

Glomm, Gerhard, Douglas Harris, and Te-Fen Lo (2005). "Charter school location," *Economics of Education Review*, 24, 451–457.

Hanushek, Eric A., John F. Kain, Steven G. Rivkin, and Gregory F. Branch (2007). "Charter School Quality and Parental Decision Making With School Choice," *Journal of Public Economics*, 91 (5/6), 823-848.

Holmes, George M., Jeff DeSimone, and Nicholas G. Rupp (2003). "Does School Choice Increase School Quality?," NBER Working Paper No. 9683.

Hoxby, Caroline M. (2004). "Achievement in Charter Schools and Regular Public Schools in the United States: Understanding the Differences," Dept. of Economics, Harvard University, mimeo.

Hoxby, Caroline M. and Jonah E. Rockoff (2004). "The Impact of Charter Schools on Student Achievement," Dept. of Economics, Harvard University, mimeo.

Hoxby, Caroline M. and Sonali Murarka (2009). "Charter Schools in New York City: Who Enrolls and How They Affect Their Students' Achievement," NBER Working Paper No.14852.

Imberman, Scott A. (forthcoming). "Achievement and Behavior of Charter Students: Drawing a More Complete Picture," *Review of Economics and Statistics*.

Imberman, Scott A. (2009). "The Effect of Charter Schools on Achievement and Behavior of Public School Students", Working paper, University of Houston.

Mazzeo, Michael (2002). "Product Choice and Oligopoly Market Structure," *RAND Journal of Economics*, vol. 33, No. 2, summer, 1-22.

Mehta, Nirav (2011). "Competition in Public School Districts: Charter School Entry, Student Sorting, and School Input Determination," Working paper, University of Western Ontario.

National Alliance for Public Charter Schools (2010). "How State Charter Laws Rank Against the New Model Public Charter School Law," January, downloaded from www.publiccharters.org.

Nechyba, Thomas J. (1999). School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers," *Journal of Public Economic Theory*, v. 1, pp. 5–50.

Nechyba, Thomas J. (2000). "Mobility, Targeting and Private School Vouchers," *American Economic Review*, 90, 130–146.

Nevo, Aviv (2000). "Mergers with Differentiated Products: The Case of the Ready-To-Eat Cereal Industry," *RAND Journal of Economics*, vol. 31, No. 3, Autumn, 395-421.

Nevo, Aviv (2001). "Measuring Market Power in the Ready-To-Eat Cereal Industry," *Econometrica*, v. 69, No. 2 (March), 307-342.

Petrin, Amil (2002). "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, vol. 110, 705-729.

Rincke, Johannes (2007). "Policy diffusion in space and time: The case of charter schools in California school districts," *Regional Science and Urban Economics*, 37, 526–541.

Sass, Tim R. (2006). "Charter Schools and Student Achievement in Florida," *Education Finance and Policy*, 1 (1), 123-138.

Seim, Katja (2006). "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, vol. 37, No. 3, Autumn, 619-640.

Su, Che-Lin and Kenneth L. Judd (2010). "Constrained Optimization Approaches to Estimation of Structural Models," Working paper, University of Chicago Booth School of Business.

Suzuki, Junichi (2010). "Land Use Regulation as a Barrier to Entry: Evidence from the Texas Lodging Industry," Working Paper, University of Toronto.

Weiher, Gregory R. and Kent L. Tedin (2002). "Does Choice Lead to Racially Distinctive Schools? Charter Schools and Household Preferences," *Journal of Policy Analysis and Management*, 21 (1), 79.

Zimmer, Ron and Richard Buddin (2003). "Academic Outcomes," in *Charter School Operations and Performance*, RAND, pp. 37-62.

TABLE 1a
Demographics and Achievement at Public, Charter and Private Schools

	All Schools			Public Schools			Charter Schools			Private Schools		
	Avg.	10th pctile.	90th pctile.	Avg.	10th pctile.	90th pctile.	Avg.	10th pctile.	90th pctile.	Avg.	10th pctile.	90th pctile.
Pct. White	14.11	0	69.77	5.60	0	22.50	2.19	0	4.00	51.01	0	85.1
Pct Black	74.47	16.04	100	81.82	36.79	100	89.03	58.00	100	38.98	8.06	99.28
Pct. Hispanic	9.18	0	26.83	10.61	0	35.09	7.98	0	26.00	5.74	0	10.89
Pct. Low Income	66.22	30.01	89.32	65.09	27.44	88.70	70.34	50.00	95.00	n/a	n/a	n/a
Reading Prof.	41.61	15.84	72.97	41.55	14.70	77.52	41.83	25.32	63.39	n/a	n/a	n/a
Math Prof.	41.68	13.51	73.98	41.54	12.99	76.10	42.18	21.05	66.78	n/a	n/a	n/a
Tract Income	\$61,500	\$28,200	\$98,400	\$56,000	\$29,300	\$136,600	\$46,000	\$25,600	\$65,600	\$92,000	\$33,341	\$142,300

Notes: The unit of observation is a campus-year. “Reading Prof.” is the percent of students who are proficient in Reading. “Tract income” is the average household income in the Census tract where the school is located. Weighted statistics; weight = Fall enrollment. Proficiency data is not available for private schools.

TABLE 1b
Demographics of Private Schools by Private School Type

	Catholic	Other Religious	Nonsectarian
Avg. Pct. White	39.34	59.03	60.99
Avg. Pct Black	49.13	32.55	29.83
Avg. Pct. Hispanic	8.47	2.51	4.52
Tract Income	\$77,100	\$105,700	\$102,200

Notes: see Table 1a.

TABLE 2
School Openings and Closings

Year	Public School Campuses			Charter School Campuses			Private School Campuses		
	Open	Closed	Merged	Open	Closed	Merged	Open	Closed	Merged
2003	0	0	0	4 (4)	0	0	0	0	0
2004	1 (1)	1 (1)	0	10 (9)	1 (1)	0	0	2	0
2005	0	5 (5)	0	9 (9)	1 (1)	0	2	0	0
2006	0	0	1 (1)	8 (7)	1 (1)	0	1	4	0
2007	0	0	1 (1)	6 (6)	3(1)	0	0	2	0

Notes: each cell indicates number of campuses; number of schools is in parentheses. For instance, in 2004 10 charter campuses and 9 new charter schools opened. The three campuses of Washington Academy were taken over by Howard Road in Spring 2008. In this table we count them as three campuses (and one school) that close in 2007.

TABLE 3a
Grade Levels at Public, Charter, and Private Schools

	Public Schools			Charter Schools			Private Schools		
	Percent (1)	Pct. of Students (2)	Avg. Enrollment (3)	Percent (4)	Pct. of Students (5)	Avg. Enrollment (6)	Percent (7)	Pct. of Students (8)	Avg. Enrollment (9)
Elementary	68.62	58.03	315	42.61	30.79	226	16.93	7.79	123
Elementary/Middle	4.28	4.96	432	20.87	23.17	347	51.12	35.43	185
Middle	14.41	15.23	394	11.74	12.56	334	0.64	0.45	39
Middle/High	n/a	n/a	n/a	6.09	6.85	352	5.75	5.29	246
High	12.70	21.78	639	14.78	19.55	413	7.99	18.75	626
Elem./Middle/High	n/a	n/a	n/a	3.91	7.08	566	17.57	32.65	496

Notes: The unit of observation is a campus-year. For instance, on average during the sample period 68.62 percent of public schools are elementary, 4.28 are elementary/middle, 14.41 are middle, and 12.70 are high schools. Among public school students, on average 58.03 percent attend elementary schools, 4.96 attend elementary/middle schools, etc. Avg. enrollment is a weighted average; weight = Fall enrollment.

TABLE 3b
Grade Levels at Different Types of Private Schools

	Catholic			Other Religion			Nonsectarian		
	Percent (1)	Pct. of Students (2)	Avg. Enrollment (3)	Percent (4)	Pct. of Students (5)	Avg. Enrollment (6)	Percent (7)	Pct. of Students (8)	Avg. Enrollment (9)
Elementary	2.27	1.71	132	32.47	17.25	147	24.04	9.50	97
Elementary/Middle	77.27	53.24	191	36.36	22.49	171	28.85	20.79	178
Middle	1.52	1.00	38.5	0	0	0	0	0	0
Middle/High	3.79	3.47	254	0	0	0	12.50	12.30	242
High	15.15	42.01	768	0	0	0	4.81	1.13	58
Elem./Middle/High	0	0	0	31.17	60.26	535	29.81	56.28	465

Notes: see Table 3a.

TABLE 4a
Demographics and Achievement By School Level

	All Schools					
	Elem.	E/M	Middle	M/H	High	E/M/H
Avg. Pct. White	8.06	17.16	3.99	40.13	12.42	54.58
Avg. Pct. Black	79.00	74.13	85.21	51.49	76.04	36.09
Avg. Pct. Hispanic	10.98	6.77	9.13	6.58	9.31	4.00
Avg. Pct. Low Income	68.96	71.14	65.71	68.99	57.61	66.17
Avg. Pct. Proficient Reading	46.18	42.44	39.95	32.89	31.89	41.01
Avg. Pct. Proficient Math	45.29	43.69	38.83	30.27	35.25	34.79
Avg. Tract Hh. Income	\$54,700	\$58,600	\$60,000	\$66,800	\$60,200	\$108,567

Notes: The unit of observation is a campus-year. "Elem." and "E/M" denote elementary and elementary/middle school, respectively. Weighted averages; weight is school-level enrollment.

TABLE 4b
Demographics and Achievement by School Type and Level

Public Schools						
	Elem.	E/M	Middle	M/H	High	E/M/H
Avg. Pct. White	6.85	1.43	4.38	0	4.07	0
Avg. Pct. Black	79.70	91.41	85.19	0	82.92	0
Avg. Pct. Hispanic	11.41	6.74	8.58	0	10.77	0
Avg. Pct. Low Income	68.32	71.31	66.66	0	53.94	0
Avg. Pct. Proficient Reading	46.86	36.92	37.21	0	31.51	0
Avg. Pct. Proficient Math	46.08	39.51	35.60	0	34.06	0
Avg. Tract Hh. Income	\$54,600	\$39,500	\$61,000	0	\$59,700	0

Charter Schools						
	Elem.	E/M	Middle	M/H	High	E/M/H
Avg. Pct. White	3.15	3.46	2.32	1.29	0.16	0.09
Avg. Pct. Black	85.20	89.25	85.52	89.24	93.97	97.37
Avg. Pct. Hispanic	10.55	6.42	11.29	8.84	5.54	1.91
Avg. Pct. Low Income	73.33	71.01	61.63	68.99	72.47	66.17
Avg. Pct. Proficient Reading	41.48	46.75	52.03	32.89	33.44	41.01
Avg. Pct. Proficient Math	39.84	46.94	53.07	30.27	40.06	34.79
Avg. Tract Hh. Income	\$50,000	\$41,600	\$55,500	\$37,700	39,000	\$53,100

Private Schools						
	Elem.	E/M	Middle	M/H	High	E/M/H
Avg. Pct. White	52.98	31.76	0	83.42	53.72	64.76
Avg. Pct. Black	41.64	58.06	58.44	9.43	35.01	24.65
Avg. Pct. Hispanic	2.49	6.98	41.56	4.07	7.39	4.39
Avg. Tract Hh. Income	\$73,300	\$76,600	\$39,600	\$99,200	\$81,000	\$118,900

TABLE 4b
Demographics and Achievement by School Type and Level (cont.)

Catholic						
	Elem.	E/M	Middle	M/H	High	E/M/H
Avg. Pct. White	0	26.48	0	82.29	53.30	0
Avg. Pct. Black	98.73	61.56	58.44	9.23	35.35	0
Avg. Pct. Hispanic	.72	9.57	41.56	4.18	7.46	0
Avg. Tract Hh. Income	\$43,100	\$74,600	\$39,600	\$74,600	\$81,600	0

Other Religious						
	Elem.	E/M	Middle	M/H	High	E/M/H
Avg. Pct. White	60.04	40.85	0	0	0	65.53
Avg. Pct. Black	35.29	50.61	0	0	0	25.02
Avg. Pct. Hispanic	1.98	1.76	0	0	0	2.95
Avg. Tract Hh. Income	\$58,400	\$82,2000	0	0	0	\$128,000

Nonsectarian						
	Elem.	E/M	Middle	M/H	High	E/M/H
Avg. Pct. White	50.94	42.87	0	83.87	75.91	64.08
Avg. Pct. Black	41.95	51.97	0	9.51	16.52	24.32
Avg. Pct. Hispanic	3.54	2.20	0	4.02	3.44	5.67
Avg. Tract Hh. Income	\$100,800	\$78,700	0	\$109,100	\$47,200	\$110,800

Notes: see Table 4a.

TABLE 5a
Program Focuses at Charter Schools

Focus	All Levels (1)	Elem. (2)	E/M (3)	Middle (4)	M/H (5)	High (6)	E/M/H (7)
Core	61.30	56.12	97.92	74.07	21.43	20.59	100
Language	11.30	22.45	2.08	11.11	0	0	0
Arts	9.57	20.41	0	0	14.29	0	0
Vocational	7.83	0	0	3.70	35.71	35.29	0
Civics	6.52	0	0	3.70	28.57	29.41	0
Math and Science	3.48	1.02	0	7.41	0	14.71	0

Notes: the unit of observation is a campus-year. “Elem.,” “E/M,” “M/H” and “E/M/H” denote elementary, elementary/middle, middle/high and elementary/middle/high school respectively. For instance, among elementary charter campuses, on average 56.12 percent focus on a core curriculum, 22.45 percent on language, etc.

TABLE 5b
Charter Schools: Student Demographics and Achievement by Level and Program Focus

a. Elementary Schools

	Arts	Core	Language	Math and Science
	(1)	(2)	(3)	(4)
Avg. Percent White	1.65	4.16	3.50	0
Avg. Percent Black	87.76	88.80	55.65	96.00
Avg. Percent Hispanic	9.12	5.82	39.56	2.00
Avg. Percent Low Income	74.51	69.20	70.34	46.00
Avg. Pct. Proficient in Reading	38.17	44.02	34.20	58.41
Avg. Pct. Proficient in Math	28.26	43.35	36.21	33.63
Avg. Tract Hh. Income	\$47,900	\$47,000	\$55,500	\$30,800

b. Middle Schools

	Civics	Core	Language	Math and Science	Vocational
	(1)	(2)	(3)	(4)	(5)
Avg. Percent White	.1	4.07	0	0	2.77
Avg. Percent Black	73.00	87.77	65.50	96.50	95.92
Percent Hispanic	25.6	6.93	34.25	2.50	.92
Avg. Percent Low Income	81.4	65.94	88.00	54	65.56
Avg. Pct. Proficient Reading	30.53	47.13	26.81	59.19	31.79
Avg. Pct. Proficient Math	32.10	46.14	21.36	57.36	30.83
Avg. Tract Hh. Income	\$44,900	\$51,300	\$57,300	\$30,800	\$36,000

c. High Schools

	Civics	Core	Math and Science	Vocational
	(1)	(2)	(3)	(4)
Avg. Percent White	.46	0.03	.30	1.03
Avg. Percent Black	76.00	97.29	99.00	97.32
Avg. Percent Hispanic	22.5	2.11	.5	1.45
Avg. Percent Low Income	72.71	66.09	100	64.93
Pct. Proficient Reading	33.95	35.22	46.29	24.64
Avg. Pct. Proficient Math	41.24	34.43	55.27	23.40
Avg. Tract Hh. Income	\$37,100	\$52,400	\$18,000	\$52,200

Note: Unit of observation is a campus-year. Weighted averages; weight = fall Enrollment. “Elementary schools” includes elementary, elementary/middle and elementary/middle/high schools; “middle schools” includes middle, elementary/middle, middle/high, and elementary/middle/high; “high schools” include high, middle/high, elementary/middle/high schools.

TABLE 6a
Types of Private School, by Level

Focus	All Levels (1)	Elem. (2)	E/M (3)	Middle (4)	M/H (5)	High (6)	E/M/H (7)
Catholic	42.31	5.66	64.15	100	27.78	80.00	0
Other Religious	24.68	47.17	17.61	0	0	0	43.64
Nonsectarian	33.01	47.17	18.24	0	72.22	20.00	56.36

Note: Unit of observation is a campus-year. "Elem.," "E/M", "M/H" and "E/M/H" denote elementary, elementary/middle, middle/high and elementary/middle/high school respectively.

TABLE 6b
Private Schools: Student Demographics by Level and Type

d. Elementary Schools

	Catholic (1)	Other Religious (2)	Nonsectarian (3)
Avg. Percent White	17.18	35.67	32.00
Avg. Percent Black	69.13	59.16	63.19
Avg. Percent Hispanic	11.66	1.63	2.59

e. Middle Schools

	Catholic (1)	Other Religious (2)	Nonsectarian (3)
Avg. Percent White	20.33	33.28	45.31
Avg. Percent Black	65.20	60.51	48.76
Avg. Percent Hispanic	12.34	1.78	3.36

f. High Schools

	Catholic (1)	Other Religious (2)	Nonsectarian (3)
Avg. Percent White	58.70	48.92	55.20
Avg. Percent Black	31.45	43.31	37.19
Avg. Percent Hispanic	6.37	2.97	4.37

Note: Unit of observation is a campus-year. Weighted averages; weight = fall enrollment.

TABLE 7
Number of Charter Campuses that Move

Year	Number of Campuses	Number of Campuses that Move
2004	39	2
2005	47	6
2006	54	7
2007	59	4

Note: unit of observation is a campus. For instance, 2 of the 39 campuses in 2004 have a different location than in 2003.

TABLE 8
Number of Campuses per Charter School

Number of Campuses per School	Number of Schools
1	35
2	4
3	5
5	1
Total	45

Note: unit of observation is a charter school organization. Since the number of campuses changes over time for some schools, the table depicts information pertaining to the highest number of campuses that each school has during the sample period. For instance, 35 charters have one campus, 4 have 2, etc.

TABLE 9
High-Performing Public and Charters Schools

	Public Schools	Charters
Avg. Tract Household Income	\$103,700	\$19,600
Avg. Percent White Students	35.56	.36
Avg. Percent Black Students	47.68	98.48
Avg. Percent Hispanic Students	10.44	.57
Avg. Percent Low Income Students	25.02	74.21

Note: unit of observation is a campus-year.

TABLE 10
Early versus Recent Charter Entrants

	Early Entrants	Recent Entrants
Number of campuses	27	37
Avg. Enrollment	450	192
Pct. Focused on Core	66.67	56.76
Pct. Elementary	14.81	59.46
Pct. Elementary/Middle	33.33	13.51
Pct. Elementary/Middle/High	11.11	0
Pct. High	22.22	2.70
Pct. Middle	11.11	18.92
Pct. Middle/High	7.41	5.41
Avg. Tract Hh. Income	\$48,800	\$50,500
Pct. belonging to multiple-campus charters	26	63
Pct. White Students	1.07	5.24
Pct. Black Students	92.54	85.05
Pct. Hispanic Students	5.94	8.55
Pct. Low Income Students	73.78	64.95
Pct. Proficient Reading	42.49	40.90
Pct. Proficient Math	40.49	35.55

Note: unit of observation is a campus. Information tabulated here corresponds to the last year that the campus is observed in the data. Weighted averages; weight is enrollment.

FIGURE 1
Number of Public, Charter and Private School Campuses

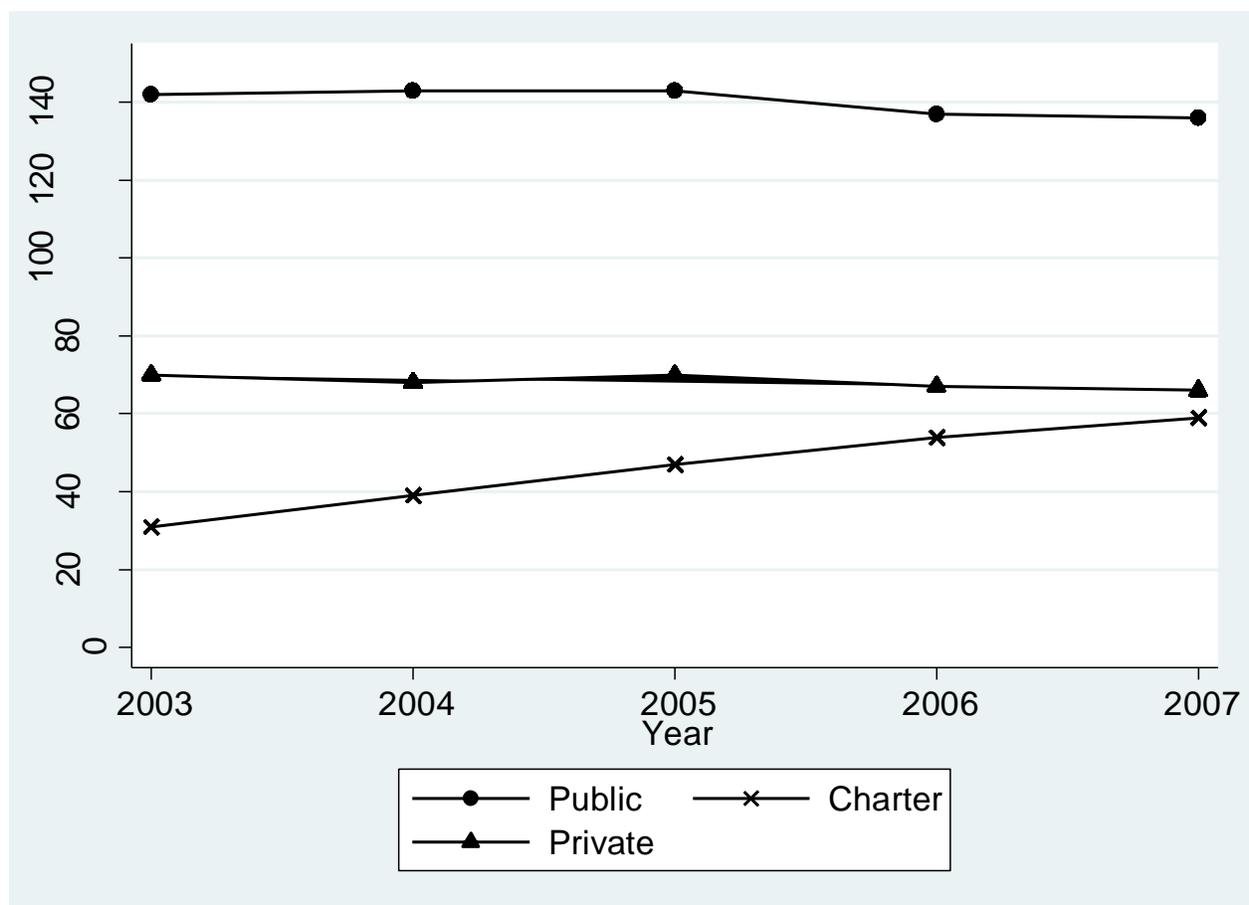


FIGURE 2
Enrollment in Public, Charter and Private School Campuses

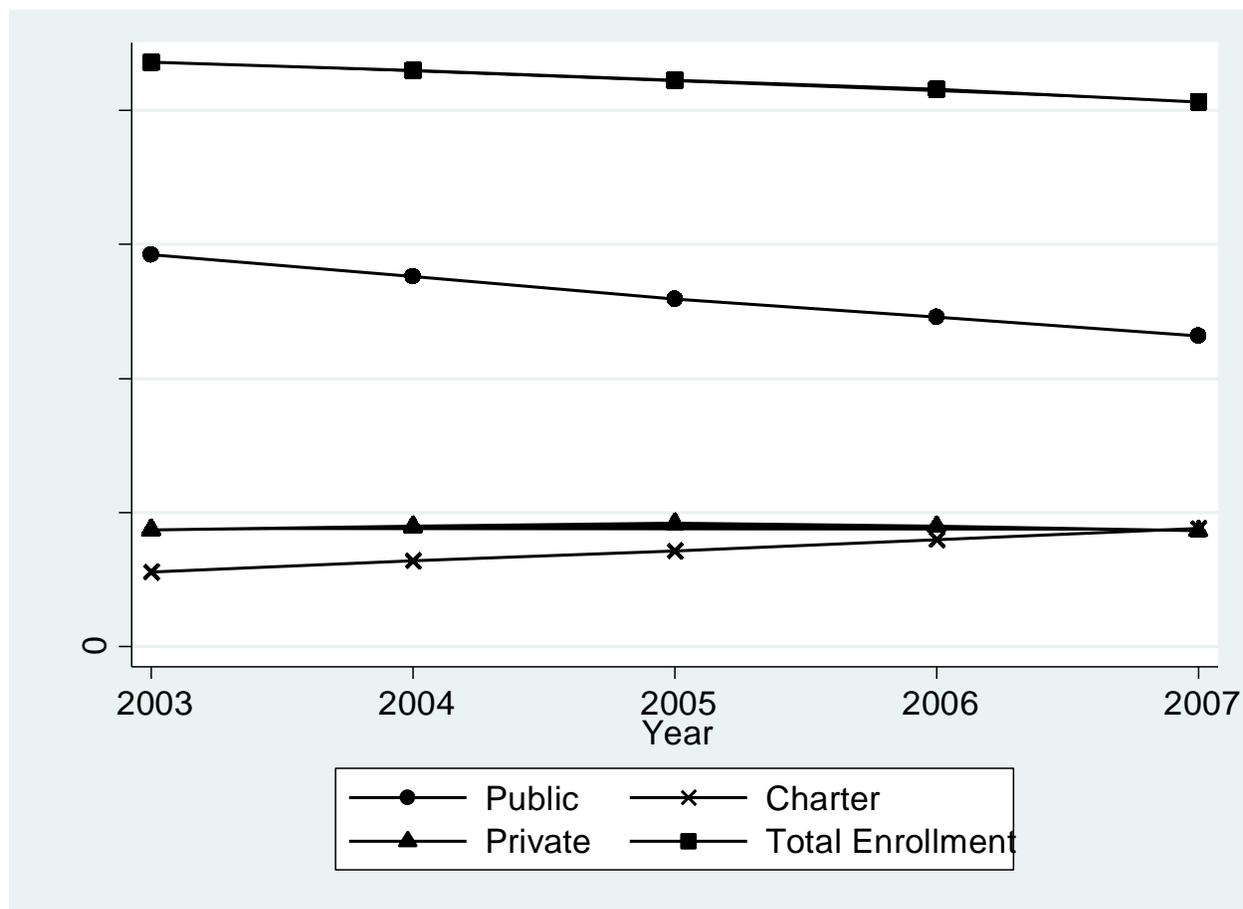
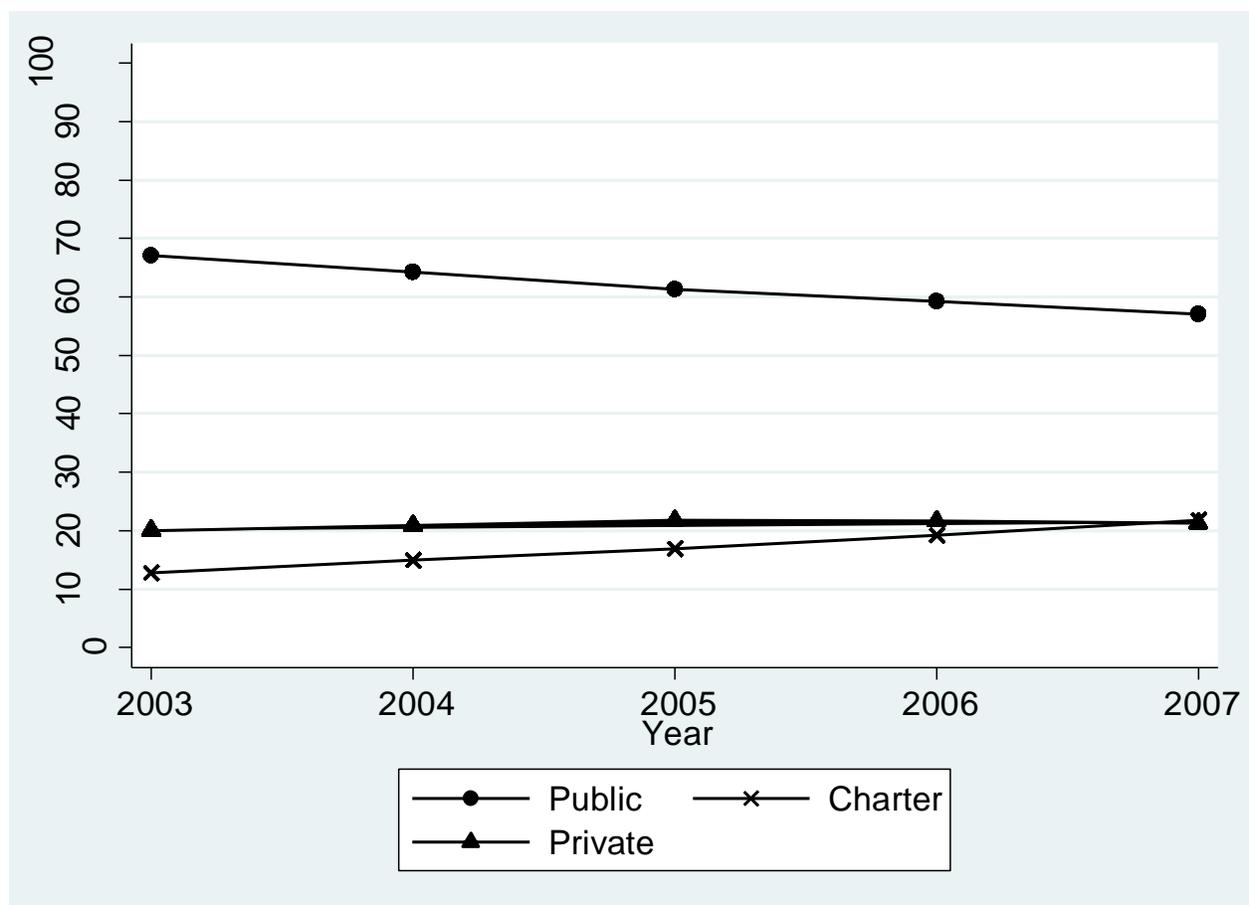


FIGURE 3
Enrollment Shares for Public, Charter and Private Schools



Notes: percentages calculated relative to total enrollment, where total is the aggregate over public, charter, and private schools for all grade levels.

FIGURE 4a
Geographic Location of Elementary Schools in DC in 2003

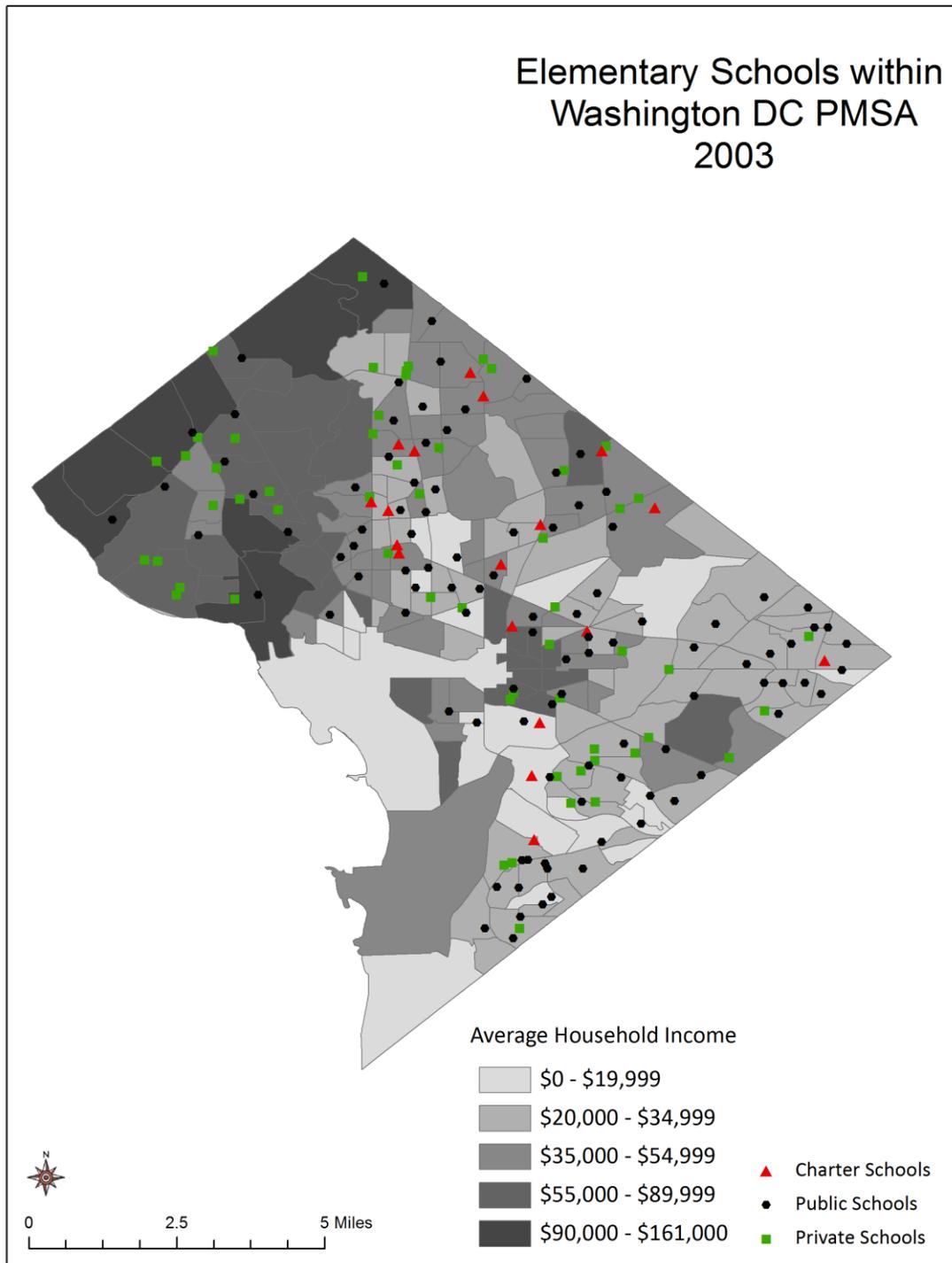


FIGURE 4b
Geographic Location of Middle Schools in DC in 2003

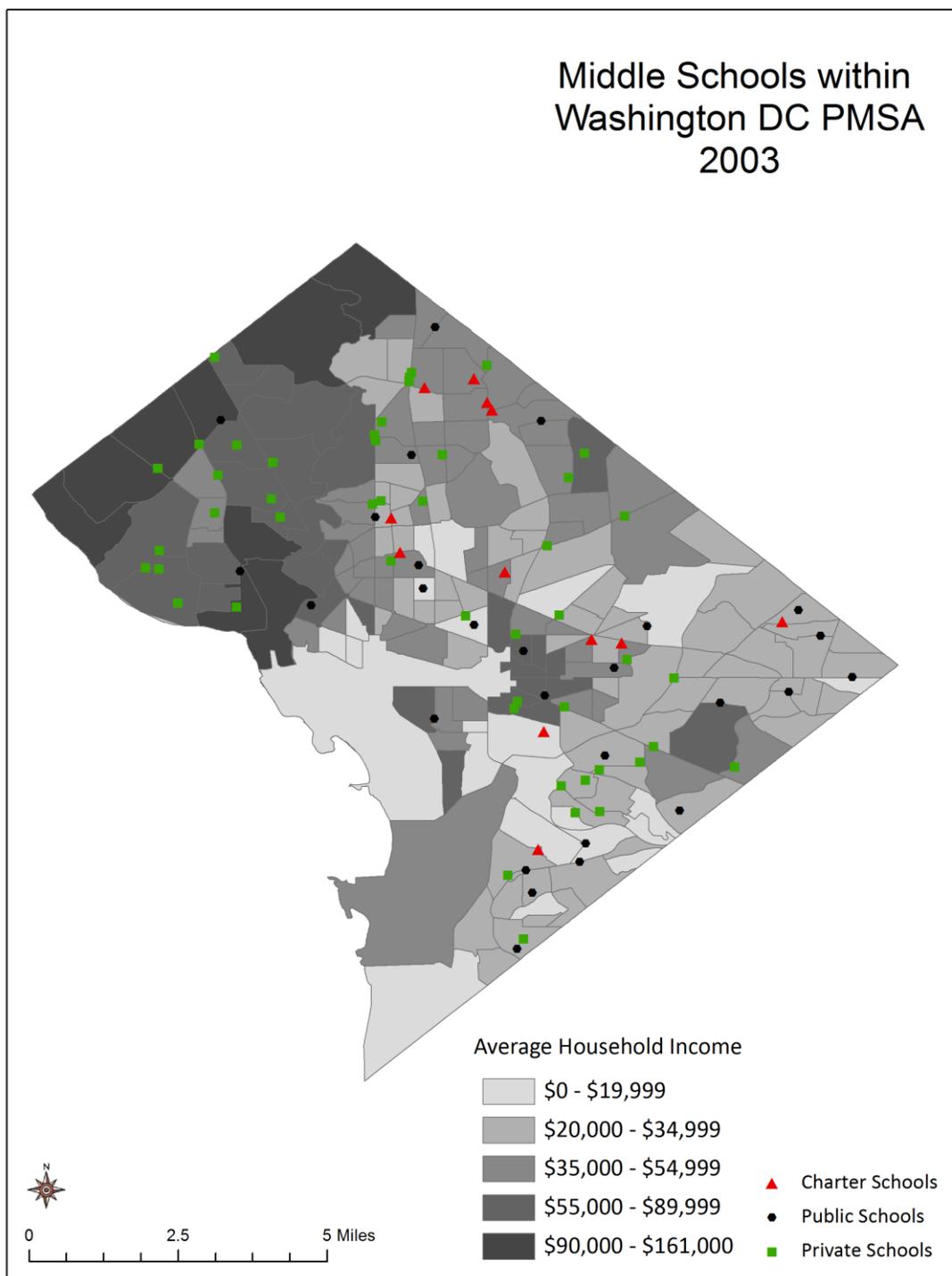


FIGURE 4c
Geographic Location of High Schools in DC in 2003

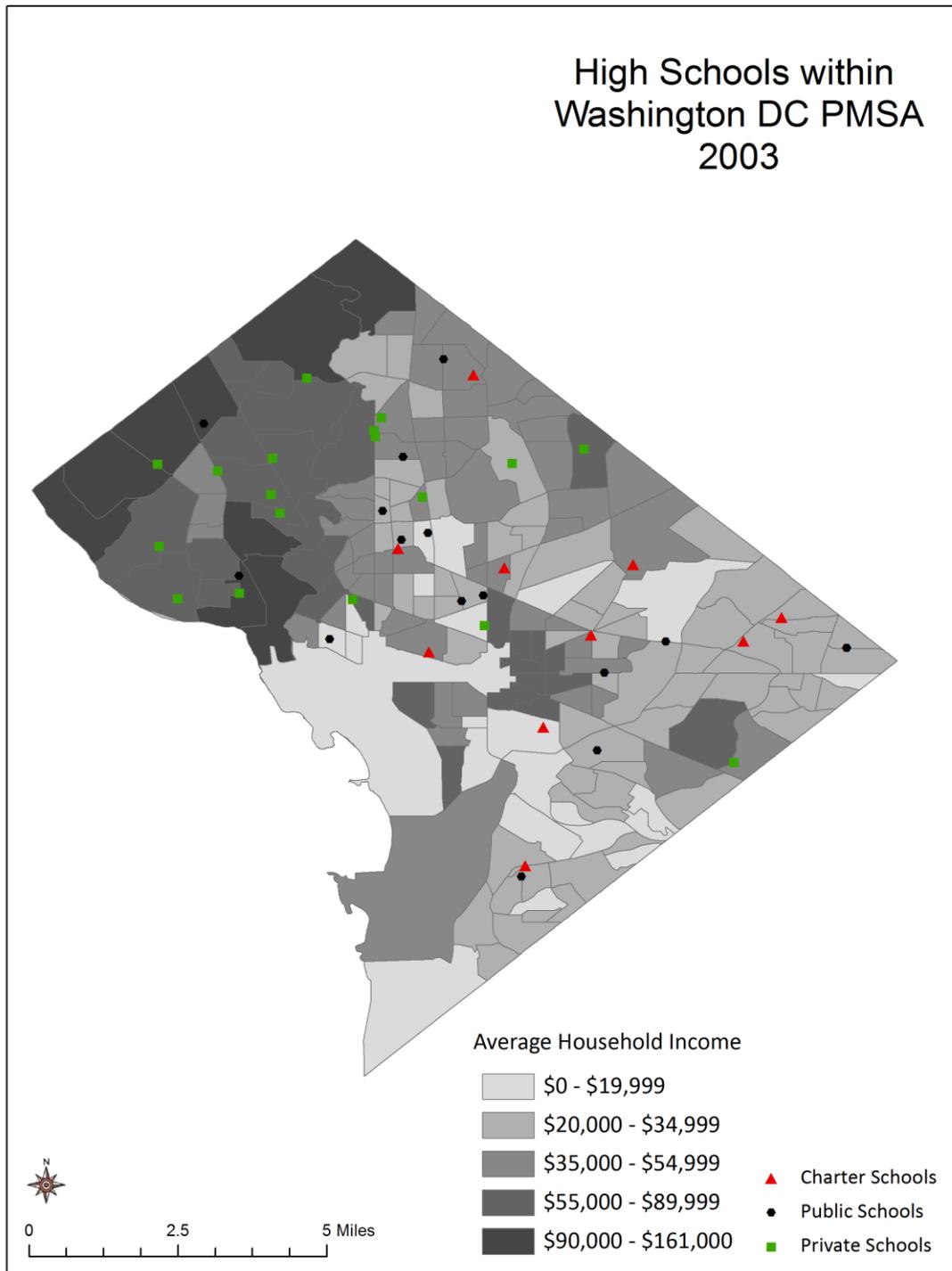


FIGURE 4d
Geographic Location of Elementary Schools in DC in 2007

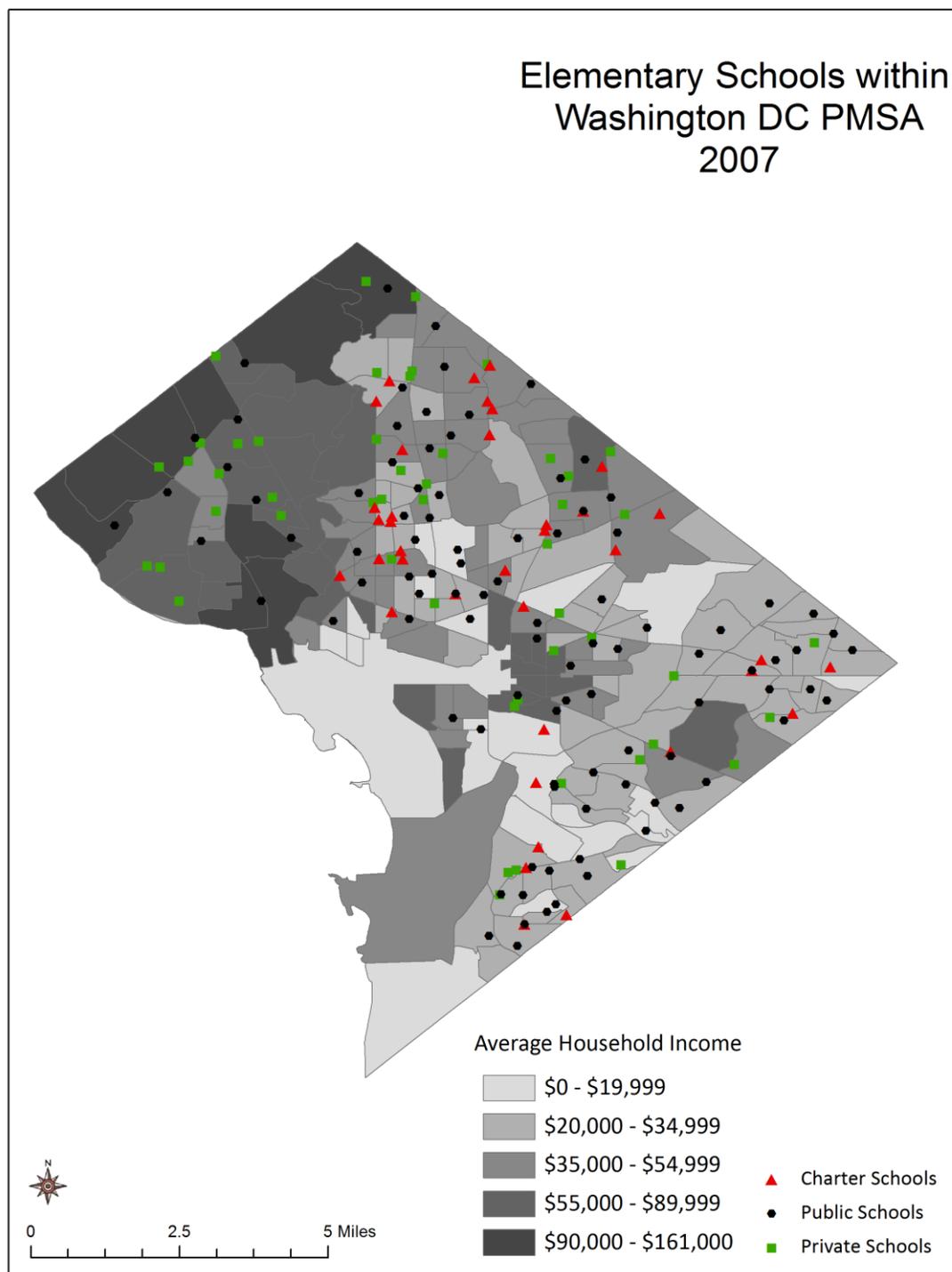


FIGURE 4e
Geographic Location of Middle Schools in DC in 2007

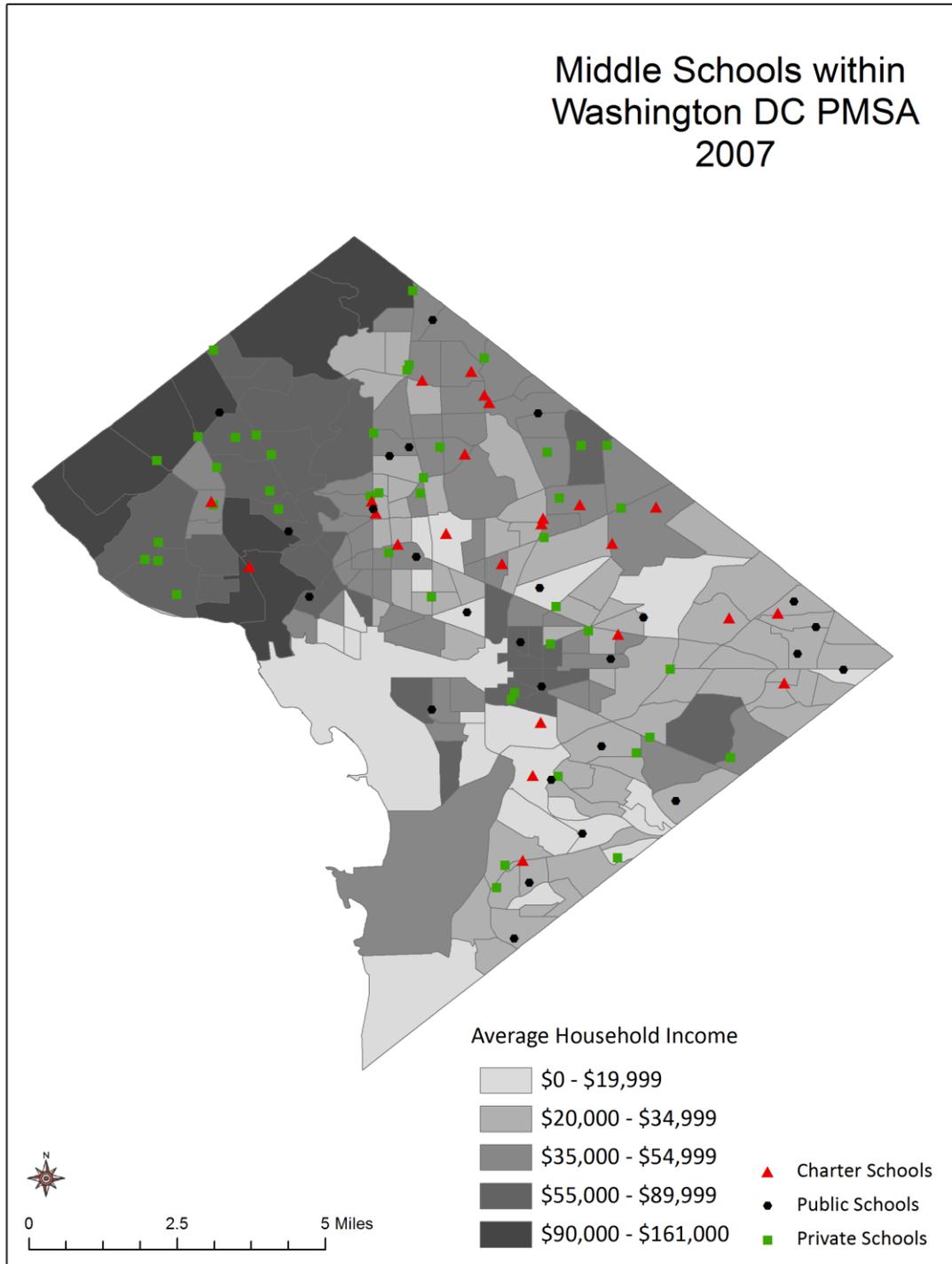


FIGURE 4f
Geographic Location of High Schools in DC in 2007

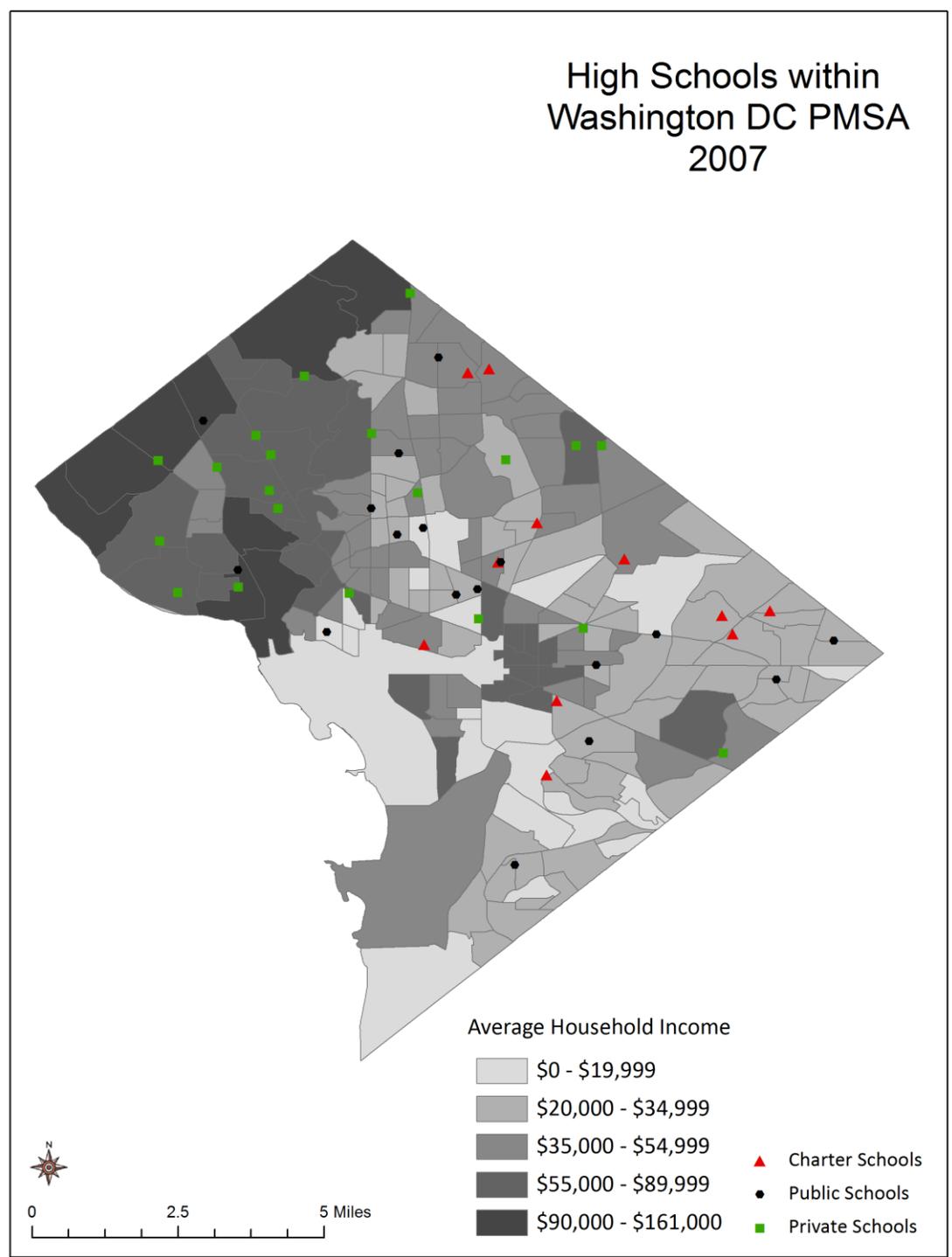


FIGURE 5a
Public Schools – Aggregate Enrollment Share by Grade

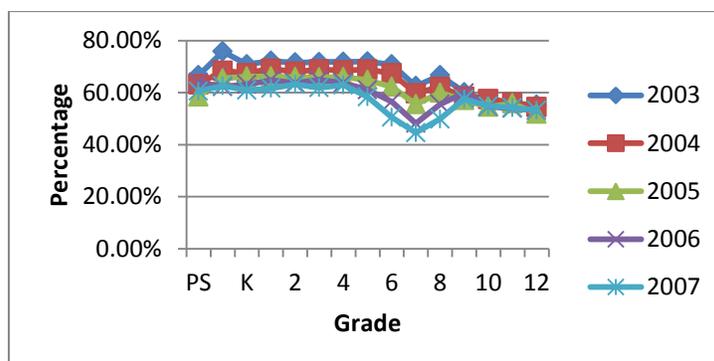


FIGURE 5b
Charter Schools – Aggregate Enrollment Share by Grade

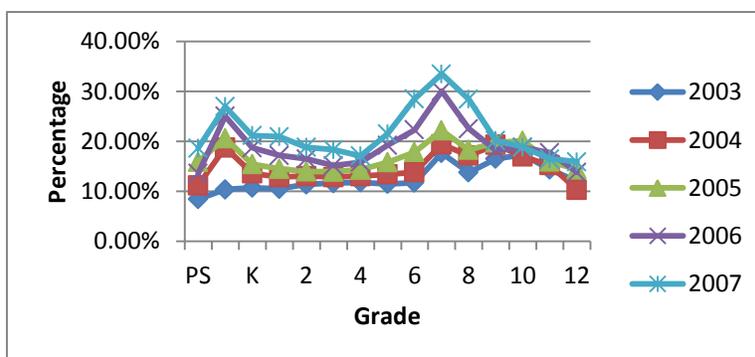
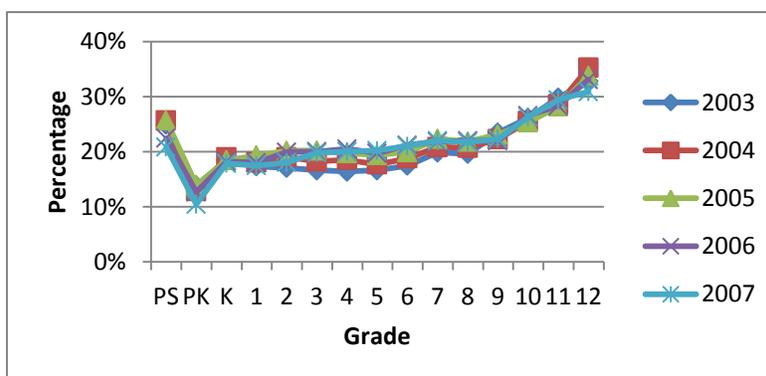


FIGURE 5c
Private Schools – Aggregate Enrollment Share by Grade



Note: shares are calculated relative to the total enrollment per grade, where total = aggregate enrollment over public, charter and private schools.

Appendix I: School-Level Data

Public Schools

The starting point for this dataset was audited enrollments from the District of Columbia Office of State Superintendent of Education (OSSE), available at <http://osse.dc.gov/seo/cwp/view,a,1222,q,552345,seoNav,31195.asp>. From here we obtained the list of public schools along with their enrollment for preschool, pre-kindergarten, kindergarten, each grade between 1 and 12, and enrollment of ungraded and adult students.

Our list only includes regular schools. This means that relative to the original OSSE list, we have excluded the following:

1. alternative schools
2. special education schools
3. early childhood centers

We excluded schools listed as alternative by the OSSE, schools that report their program as alternative in the Common Core of Data (CCD), and schools listed as alternative during the corresponding years by DCPS. We also excluded schools that offer residential programs, such as Ballou Stay and Spingarn Stay. In total, we eliminated 11 alternative schools from the original OSSE lists.

We excluded schools listed as special ed by OSSE, and schools that report their own type as special ed in the CCD. Some schools report their type as special ed only during one or two years during our sample period, in which case we called the individual schools to gauge their type and removed them from the data if they were special ed. In total, we eliminated 22 special ed schools from the original OSSE lists.

Early childhood centers include one or several of the following: pre-school, pre-kindergarten, and kindergarten, and do not include regular grades. We excluded early childhood centers as long as they never had enrollment in regular grades during the sample period. For instance, if an early childhood center added first grade in 07, then the school was included for all 5 years of our sample.

Below is the list of information and sources for each school, along with the construction of the corresponding variables:

Address: geographic address of the school. The main source for this variable was the CCD. When the address was not consistent over time, we used Google Earth and Google Maps to see whether the different addresses corresponded to the same geographic location or not. If not, we consulted the school's web site, or called the school to track down the history of its location. In the absence of other information, we used the most recently reported address. We geocoded all addresses.

School enrollment: total school enrollment, excluding ungraded and adult students. Source: own calculations based on OSSE.

Grade-level enrollment (grades preschool through 12). Source: OSSE.

Percent of white students: percent of White students in the school. Source: own calculations based on CCD. For the few cases in which the CCD data was not available, we used the demographics reported to fulfill NCLB requirements (found in <http://www.nclb.osse.dc.gov/>). Notice, however, that the NCLB requirements pertain to the students enrolled in the grades tested for the sake of NCLB accountability, not to the entire student body. When using the NCLB web site, we computed this percent as the ratio between the number of white students in the grades tested and the total number of students in those grades.

Percent of black students, percent of Hispanic students, percent of students of other ethnicities: constructed similarly to percent of white students.

Percent of Low Income students: calculated as the percent of students who receive either free or reduced lunch. Source: own calculations based on CCD. For the few cases in which the CCD data was not available, we used the demographics reported in fulfillment of NCLB requirements.

Reading proficiency: percent of students who are proficient in Reading. Source: <http://www.nclb.osse.dc.gov/>. In 03 and 04, proficiency levels were determined according to the Stanford-9 assessment. To be considered proficient, a student was supposed to score at the national 40th percentile or higher. Since 05, proficiency has been determined according to DC CAS (Comprehensive Assessment System). See <http://www.nclb.osse.dc.gov/faq.asp> for more information.

Prior to 05, the grades tested were 3, 5, 8 and 10 (according to the School Performance Reports for PCBS-authorized charter schools, and according to our own calculations comparing grade-level enrollment with number of students tested). Since 05, the grades tested have been 3, 4, 5, 6, 7, 8 and 10 (according to <http://www.nclb.osse.dc.gov/aboutayp.asp>, School Performance Reports and our own calculations comparing grade-level enrollment with number of students tested).

The achievement website, <http://www.nclb.osse.dc.gov/faq.asp>, reports proficiency for elementary and secondary schools. According to <http://www.nclb.osse.dc.gov/faq.asp>, schools are classified as elementary or secondary as follows. Since 05, elementary schools are schools with a 3rd and/or 5th grade. If the school goes above the 6th grade it must also have a 3rd grade. Secondary schools are schools with no 3rd and a grade above 6th. Prior to 2006, schools with a wide grade range were required to meet both the elementary and secondary targets. The general rule for most schools was that they were classified as elementary if they had no grade above 6 and as secondary if they had no

grade below 5. Schools that did not meet either of these criteria had to meet the targets at both levels.

For some schools and years, proficiency data are not available for one of the following three reasons: 1) the school only includes early childhood enrollment; 2) the school only includes grades that are not tested (for instance, the school includes kindergarten and first grade only); 3) the school includes grades that are tested, but enrollment in those grades is below the minimum threshold for reporting requirements. The last reason is the most prevalent cause of missing proficiency. See below for the imputations made in those cases.

Math proficiency: percent of students who are proficient in Math. Constructed similarly as reading proficiency.

Percent of out-of-boundary enrollment: percent of students enrolled through the out-of-boundary process for years 04, 05, 06 and 07 (we do not have information for 03). Source: own calculations based on DCPS data.

Year of Opening: year the school opened, if it was not open in 03. Using the CCD “status” variable and web searches we verified the school’s initial year. This is the first year for which we have records.

Year of Closing: year the school closed, if it was not open in 07. The variable stores the first year that the school was no longer open. We verified the content of this variable using the CCD “status” variable and web searches.

Year of Merge: year the school merged with another school. The variable stores the first year that the school no longer operates separately, which is the first year for which we have joint records.

Ethnic composition and low-income status of the student body were missing for 2 and 4 (out of 701) observations respectively; these are schools with very low enrollment. Since ethnic composition varies very little for any given school over time, to the cases of missing ethnic composition we imputed the school’s average ethnic composition over the years for which we do have data. Whenever possible, we imputed the predicted value coming from a school-specific linear trend.

Achievement was missing in 16 out of 701 observations. To these observations, we imputed the predicted achievement coming from the regression of school-level proficiency rates on year dummies, ethnic composition variables, percent of low income students, enrollment, and school fixed effects. In cases in which we had no proficiency data for the school at all, we ran a similar regression excluding school fixed effects and including dummies for school level, and used the resulting predicted values for our imputations.

Charter Schools

As with public schools, the starting point for this dataset was audited enrollments from the District of Columbia Office of State Superintendent of Education (OSSE), available at [http://osse.dc.gov/seo/cwp/view,a,1222,q,552345,seoNav,\[31195\].asp](http://osse.dc.gov/seo/cwp/view,a,1222,q,552345,seoNav,[31195].asp). From here we obtained the list of public schools along with their enrollment for preschool, pre-kindergarten, kindergarten, each grade between 1 and 12, and enrollment of ungraded and adult students.

For consistency with public schools, we excluded alternative schools. We identified alternative schools based on OSSE's classification, the schools' mission statements, and the self-reported program type in the CCD. We excluded schools in which ungraded or adult students constitute the majority of the student body, and schools with residential programs.

By law, charter schools cannot serve special education students exclusively. However, they can offer services targeted to specific populations even though their enrollment must be open to all students. Only one charter school in D.C., St. Coletta, serves special ed students exclusively, and they made this agreement with the Board of Education (BOE) when they were chartered. We excluded schools whose services target mostly special ed students. We identified those schools based on OSSE's classification, the schools' mission statements, the self-reported program type in the CCD, and phone conversations with staff from the Public Charter School Board (PCSB).

We excluded early childhood schools only if they never added regular grades. We also excluded online campuses.

Some non-early childhood charters opened an early childhood campus during the sample period. We only included these campuses if at some point they added regular grades. Community Academy-Amos II is an early childhood campus that did not include regular grades (at least not during our sample period) and is hence not included in our data. Similarly, Roots Academy opened an early childhood campus in 07, which we did not include in the data. KIPP-LEAP is an early childhood campus that opened in 07 and is not included either. In contrast, E.L. Haynes opened a campus in 07 for early childhood and first grade; this campus is included in our data.

Below is the list of variables for charter schools:

Address: geographic location of the campus. For PCSB schools, the main source was the SPRs. For BOE schools, the main source was the CCD. We supplemented these sources with web and Internet archive searches, and phone calls to the individual schools. We geocoded all addresses.

Several schools moved in the middle of the school year, temporarily relocated, or closed. We consulted the SPRs and various web sites to handle these cases. If the school moved in the middle of a school year, the address variable contains the address for the new location. Some schools relocated some students for a few months during renovations.

Since this was a temporary arrangement and parents knew it, we did not consider this a change in address.

School enrollment: total school enrollment, excluding ungraded and adult students. Source: own calculations based on OSSE.

Grade-level enrollment (grades preschool through 12). Source: OSSE.

Percent of white students: percent of White students in the school. For PCSB charters, the source is the School Performance Reports (SPRs) when available. Only for one campus and year (Tri Community in 07), does the SPR not include demographic information, in which case we use the NCLB data from <http://www.nclb.osse.dc.gov/> otherwise.

For BOE charters in 2007, the source is the SPRs (in 2007, the PCSB began including the BOE-authorized charters in its reports). For BOE charters before 2007, the source is the NCLB web site. Note that the NCLB information pertains only to students in grades tested, not to the whole student body of a school. These sources were supplemented by the CCD when needed.

When the school has multiple campuses but we only have one set of ethnic composition data, we impute it to all campuses.

Percent of black students, percent of Hispanic students, percent of students of other ethnicities: constructed similarly to percent of white students.

Percent of Low Income students: percent of low-income students, equal to the fraction of students on free or reduced lunch. For PCSB charters, the source is the SPRs. For BOE charters, the source is NCLB information in <http://www.nclb.osse.dc.gov/>. These sources were supplemented by the CCD when needed and possible, in which case case pctLowInc is the fraction of students with either free or reducedlunch.

When the school has multiple campuses but we only have one set of low-income variables, we impute it to all campuses.

Reading proficiency: percent of students who are proficient in Reading. See public schools for sources and construction.

Some charters span elementary and secondary grades, and this affected their reporting. For instance, Capital City had to meet elementary and secondary targets before 05, and for those years it reported two sets of scores. Since 05, Capital City has had to meet only one target, has been considered an elementary school for the sake of reporting, and has had to report only one set of scores. In general, when the school had only one campus but separate proficiencies for elementary and secondary students, we combined them into a

single proficiency indicator for comparability with other years in which we had a single proficiency indicator.

For multi-campus charters we usually had achievement data for each campus. When we did not, we imputed the available data to all the campuses.

As with public schools, we did not have proficiency data for some campus and years for the reasons enumerated above. In the case of many PCSB schools for which the NCLB web site did not report test scores due to low enrollment, we obtained proficiency rates from the SPRs. This was not possible for BOE schools with low enrollment since the SPRs only cover PCSB schools before 07. In cases in which we could not find proficiency data, we made imputations (see below).

Math proficiency: percent of students who are proficient in Math. Constructed similarly as reading proficiency.

Year of Opening: year the campus opened. Source: SPRs, FOCUS, web searches. The variable stores the first Fall that the school is open.

Year of Closing: year the school (as opposed to the campus) closed, even if it closed after 07. The variable stores the first year that the school is no longer open. The sources are Center for Education Reform (http://www.edreform.com/accountability/charters/CER_ClosedCharterSchools2009.pdf), current SPRs and PCSB listings of charter schools, current NCLB reports, and web searches.

Note that Washington Academy was taken over by Howard Road in Spring 08 – i.e., in the middle of school year 07. Hence, the year of closing is 2007 for Washington Academy because it the first year for which Washington Academy no longer exists. In fact, for the school year 07/08, data were reported by Howard Road, and the SPRs report data for Howard Road campuses noting that they were former campuses of Washington Academy. We chose not to add campuses to Howard Road for 07 because the campuses did not physically move – they simply changed ownership.

Reason for Closing: reason for closing (academic / financial / mismanagement). Source: Center for Education Reform, SPRs, web searches.

Statement: the school's mission statement. Source: schools' web sites, FOCUS, SPRs.

The percent of low-income students was missing for 9 out of 230 observations respectively. As in the case of public schools, these schools had low enrollments. In most cases we missing percent of low-income students we imputed a value equal to the school's average percent of low-income students, the average calculated over the years for which the school does have data.

In the case of missing proficiency rates (36 out of 230 observations), we made imputations similar to those described for public schools, the only difference being that we used school- (as opposed to campus-) fixed effects in the regressions.

Private Schools

The starting point for this dataset was the list of private schools from the Private School Survey (PSS). Since the PSS is biennial, we included the 2003, 2005 and 2007 waves in our data.

PSS classifies schools as regular, vocational, special ed, and other/alternative. 92 percent of the schools in our dataset are regular, and the remaining schools are classified as other/alternative. Although an alternative public school is usually a school for students with behavioral problems, an alternative private school is often a school that offers a specialized rather than a regular curriculum. In these cases, we kept alternative schools. We eliminated vocational schools because they enrolled exclusively ungraded students. We also eliminated special ed schools, early childhood centers (as long as they never had enrollment in regular grades during the sample period), and schools that only taught ungraded students. Most of the schools we eliminated are early childhood centers.

Since 2004 was not included in PSS, we assigned 2004 values to the variables through linear interpolation of 2003 and 2005, and similarly for 2006. For instance, we calculated the percent of White students for a school in 2004 as the average of percent white in 2003 and 2005.

Some schools did not report to the survey for some waves. For instance, Edmund Burke has data for 2001, 2005 and 2007 but not for 2003, which means that we need to make some data imputations for 2003, 2004 and 2006. The imputed data for 2003 is the average between 2001 and 2005. Treating the imputed 2003 as actual data, we then imputed data for 2004 as the average between 2003 and 2004.

If a school does not appear again in PSS after a particular wave, we assumed that the last year of operation was the year of the last wave in our data. Similarly, if a school appears in a particular wave but not in any of the following waves, we assume that its last year is the year of the last wave for which we have data.

Below is the list of information and sources for each campus, along with the construction of the corresponding variables:

Address: geographic location of the campus. We supplemented PSS with web and Internet archive searches, and phone calls to the individual schools. We geocoded all addresses.

School enrollment: total school enrollment, excluding ungraded and adult students.
Source: own calculations based on PSS.

Grade-level enrollment (grades pre-kindergarten through 12). Note that the PSS questionnaire pools preschool and prekindergarten in one grade level. Hence, we split the enrollment from this category equally between preschool and prekindergarten.

Percent of white students: percent of White students in the school. Source: own calculations based on the reported number of White students and the total enrolment. Since the number of students for each ethnicity spans grades K through 12, we used K-12 enrollment as well in the denominator.

Percent of black students, percent of Hispanic students, percent of students of other ethnicities: constructed similarly to percent of white students.

Tuition: annual tuition by grade. These tuition correspond to the 2010/2011 school year and were obtained from web sites and phone calls to the individual schools. For the cases in which we did not obtain tuition data, we imputed tuition based on the predicted values arising from the regression of tuition on grade, school type (based on the nine typologies used by PSS), average household income of the tract where the school is located, school enrollment, a quadratic term in school enrollment. Both observed and predicted tuitions are expressed in dollars of 2000.

Appendix II: Achievement Regression

Recall from (6) that achievement at the individual level is defined as follows:

$$1. A_{ijgt} = y_j \beta'' + \bar{D}_{jt} \alpha'' + D_i \omega'' + y_j D_i \tilde{\beta}'' + \xi_{jgt}'' + v_{ijgt}$$

In this equation, the productivity shock ξ_{jgt}'' is decomposed as follows:

$$2. \xi_{jgt}'' = \xi_j'' + \xi_g'' + \xi_t'' + \Delta \xi_{jgt}''$$

In order to derive mean achievement in school j at time t , for simplicity we assume a finite number of students in each school and grade. Recall that N_{jgt} is the number of students in school j 's grade g at time t , and define N_{jt} as the number of students in school j at time t . Recall that the school's level κ_j is the set of grades encompassed by the school. Note that school j 's demographic composition at time t , \bar{D}_{jt} , is the average of the demographic characteristics D_i of

the students who choose to attend school j : $\bar{D}_{jt} = \frac{\sum_{g \in \kappa_j} \sum_{i \in jgt} D_i}{N_{jt}}$. Thus, average achievement in

school j at time t is equal to the following:

$$3. A_{jt} = y_j \beta'' + \bar{D}_{jt} \rho + y_j \bar{D}_{jt} \tilde{\beta}'' + \frac{N_{jt} \xi_j''}{N_{jt}} + \frac{\sum_{g \in \kappa_j} N_{jgt} \xi_g''}{N_{jt}} + \frac{N_{jt} \xi_t''}{N_{jt}} + \frac{\sum_{g \in \kappa_j} \sum_{i \in jgt} v_{ijgt}}{N_{jt}} + \frac{\sum_{g \in \kappa_j} N_{jgt} (\Delta \xi_{jgt}'')}{N_{jt}}$$

where $\rho = (\alpha'' + \omega'')$. We can rewrite this equation as follows:

$$4. A_{jt} = y_j \beta'' + \bar{D}_{jt} \rho + y_j \bar{D}_{jt} \tilde{\beta}'' + \bar{\xi}_j + \bar{\xi}_t + \bar{v}_{jt}$$

Here, $\bar{\xi}_j = \xi_j'' + \frac{\sum_{g \in \kappa_j} N_{jgt} \xi_g''}{N_{jt}}$ and $\bar{\xi}_t = \xi_t''$. Although the second term of $\bar{\xi}_j$ is not strictly constant

given that the contribution of grade g to school j 's enrollment potentially varies from year to year, in most of our sample this contribution varies very little over time. Thus, it can be reasonably approximated by a grade-specific constant, which makes the second term of $\bar{\xi}_j$ a school-specific constant. As a result, $\bar{\xi}_j$ embeds both a school fixed effect and a school-specific weighted average of grade-level fixed effects. The time fixed effect $\bar{\xi}_t$ captures changes in

achievement that affect all schools and grades (such as the change in assessment instrument that took place in 2006). Finally, the error term \bar{v}_{jt} is defined as follows:

$$\bar{v}_{jt} = \frac{\sum_{g \in \mathcal{K}_j} \sum_{i \in jgt} v_{ijgt}}{N_{jt}} + \frac{\sum_{g \in \mathcal{K}_j} N_{jgt} (\Delta \xi_{jgt}^v)}{N_{jt}}.$$

In other words, the error term embeds the weighted mean of the school's grade-specific idiosyncratic productivity shocks and the mean of the idiosyncratic components of performance of the school's students. Since parents observe ξ_{jgt}^v when choosing schools, the school's demographic composition \bar{D}_{jt} is correlated with \bar{v}_{jt} . Thus, the achievement regression must be estimated using instrumental variables.

Appendix III: Market Size and Outside Good

Since each grade is a market, calculating market size amounts to determining the number of children who are eligible for each grade in Washington, D.C. in each year between 2003 and 2007. Given our data on aggregate enrollment by grade and year, the share of the outside good for a given grade and year is equal to (market size – aggregate enrollment) / market size.

The main difficulty in calculating market size is that we do not have direct information on the number of children eligible for each grade in each year. In the absence of this information, the count of children by age and year would be helpful, but we do not have this information either. Instead, to calculate market size we rely on the following pieces of data:

1. The 2000 count of children by age (see Table A1);
2. The intercensal estimates of the number of children in the below 5, 5-13 and 14-17 year old brackets (see Table A2);
3. The 2000 count of enrolled and not enrolled children, and the resulting percent of children who are not enrolled. The latter is our best proxy for the outside good share for year 2000;
4. Observed enrollment for each grade and year (see Table A4).

To estimate the number of children by age in each year, one could apply the annual rates of growth implied by Table A2 to the 2000 observed number of children in each age (see Table A1). Since these rates of growth differ by age groups, one could apply to each age the rate that corresponds to its age group. In principle, the resulting number of children per age should be at least as large as the observed enrollment in the corresponding grades (see Table A4). When we apply this procedure, for grades 4-8 we estimate a number of children (or potential enrollment) that is *lower* than the actual enrollment, thus implying a negative share of the outside good. One possible explanation is the existence of different growth rates within the age group 5-13. For instance, children who are 5-13 years old in 2003 would have been 2-10 years old in 2000. As Table 1 shows, there is a fair amount of variation in the number of children in each age between 2 and 10. Thus, it is plausible that each individual age would grow at a different annual rate. Another possible explanation is the existence of complex patterns of grade retention (i.e., children who must repeat a grade).

Since we do not have data to disentangle these possibilities, after evaluating a number of solutions we opted for inflating the observed enrollment for each grade and year by a factor \mathcal{G}_{gt} . These factors are chosen so that their implied outside good shares match up with 2000 Census outside good shares (see Table A3). We opted for this solution not only because it was simple, but also because our most direct evidence on the outside good share comes from the 2000 Census. Thus, we calculated market size for (g, t) as observed aggregate enrollment for (g, t) times \mathcal{G}_{gt} .

For the calculation of our adjustment factor, we modified the 2000 Census outside good shares slightly to accommodate some features of our data. In particular, our enrollment data is based on regular schools (thus excluding special ed or alternative schools, which are more prevalent in

middle or high than elementary school), and excludes schools that are specialized in early childhood (i.e., that do not include grades above kindergarten). Thus, we used an outside good share of 70% for preschool and 45% for prekindergarten; 3% for grades K through 8 (corresponding to ages 5-14 in the 2000 Census), and 10% for grades 9 through 12 (roughly corresponding to ages 14-17 in the Census).

Our solution has the appealing feature of delivering an estimated number of children whose annual growth rate is consistent with the growth rates implied by the intercensal Census estimates (see Table A2). In particular, potential enrollment given our solution grows at the following rates between 2003 and 2007: 4% for preschool and prekindergarten, -13% for grades K through 8, and 7% for grades 9 through 12. These rates line up with the Census growth rates for the corresponding age groups (equal to 7%, -13% and 13%, respectively) for the same period.

Appendix IV: Household Types and Choice Sets

In this appendix we first describe the construction of household types for year 2000 and the adjustments needed for years 2003-2007. Then we describe the construction of choice sets for households in different locations.

Household Types for Year 2000

The main challenge in measuring the number of households in each (grade, demographic type) combination at the block group level is the fact that we do not observe the joint distribution of child age, race, household income and poverty status at the block group level. Instead, the Census provides us with the following information:

- a. tract-level joint distribution of age and race;
- b. tract-level joint distribution of age bracket, race and poverty status;²⁸
- c. tract-level joint distribution of family income (by brackets) and race;
- d. block group-level joint distribution of age brackets and race.

Recall that Washington, D.C. includes 433 block groups and 188 Census tracts. The calculations described below apply to the 185 tracts that do have children aged 0-18 and the corresponding block groups.

We proceeded as follows:

1. Note that for each block group we observe the joint distribution of age brackets and race, and at the tract level we observe the joint distribution of *individual* ages and race. We assume that for a given age bracket and race, the distribution of individual ages is the same for all the block groups that belong to a tract. Hence, for each block group, age bracket and race we imputed the distribution of individual ages of the corresponding tract and race.
2. For each tract we observe the joint distribution of age brackets, race and poverty status. We assume that for a given tract, the distribution of race and poverty status is the same for all individual ages within a given age bracket. We also assume that the joint

²⁸ Federal guidelines (see <http://www.fns.usda.gov/cnd/governance/notices/iegs/iegs.htm>) establish eligibility for free- or reduced-price lunch based on household size and income. To qualify for free (reduced) lunch, a child must live in a household whose income is below 130 (185) percent of the Federal poverty guidelines for that household size. We pool children eligible either for free or reduced lunch into a single category. Since the Census reports the number of children in each age bracket and race whose household income falls below 130%, or between 130% and 185% percent of the poverty guideline given their household size, we can calculate the number of children eligible for free- or reduced-lunch as the total number of children whose household income falls below 185% of the poverty guideline. Thus, “poverty status” is a binary variable that describes whether the child is eligible for free- or reduced-lunch or not. In some parts of the paper, we use “low-income” to describe a child who is eligible for free- or reduced lunch.

distribution of age, race, and poverty status is the same for all the block groups located within a given tract. Hence, for each block group, individual age and race, we imputed the poverty distribution of the corresponding tract.

3. Based on steps (1) and (2), we calculated the number of children in each (age, race, poverty status) combination in each block group. For a given block group, denote this number as n_{arp} , where a =age; r =race; p =poverty status.
4. For each tract we observe the joint distribution of family income and race, but not the joint distribution of family income, race and poverty status. To determine how many of the n_{arp} children in the corresponding (a, r, p) combination a given block group fall in each income bracket, we perform an imputation which we can illustrate through the following example for a hypothetical block group and tract. In this block group, 20% of 5-year old White children are poor, and the remaining 80% are not. In the corresponding tract, 5% of White families have incomes below \$20,000; 15% have incomes between \$20,000 and \$40,000, and the remaining 80% have incomes between \$40,000 and \$60,000. Thus, we assigned a family income of \$10,000 (i.e., the midpoint for the income bracket between \$0 and \$20,000) to a quarter of the 5-year old White children, where $1/4 = 5\% / 20\%$. Similarly, we assigned a family income of \$30,000 (i.e., the midpoint for the income bracket between \$20,000 and \$40,000) to three-quarters of the 5-year old, White children, where $3/4 = 15\% / 20\%$. To the 80% of 5-year old White children who are not poor, we assigned an income of \$50,000 (i.e., the midpoint of the income bracket between \$40,000 and \$60,000).
5. Based on step (4), for each block group we calculated the number $\mu_{\ell am}$ of children of each age and demographic type m , where the demographic type is given by a (race, income, poverty status) combination.
6. To calculate the number of children of each grade g and demographic type m , $\mu_{\ell gm}$, we calculated the block-group average of the two most frequent ages in the grade (with the exception of preschool, as explained below), given the lack of data on the age distribution of children by grade. For instance, for location ℓ and demographic type m , we calculated the number of children who are eligible for preschool as the number of 3-year olds; the number of children who are eligible for prekindergarten as the average of the number of 4- and 5-year olds, and so on until 12th grade, for which we took the average of 17- and 18-year old children.

Calculating the number of children aged 18-years old in each demographic type and location is challenging for D.C. because the number of 18-year olds is much higher than the number of 17-year olds, and the demographics of these two ages are quite different as well. Similarly, the age bracket 18-24 has different demographics than the age bracket 12-17. We believe this is because many 18 year olds in D.C. attend college and do not come from D.C. Hence, we determined the number of 18-year olds at the block group level as the average number of children by age in the 12-17 year old bracket. We assigned to 18-year olds the same demographics as the average of the 12-17 year old bracket.

Household Types for 2003-2007

We obtained the market size for each grade and year as explained in Appendix III. This gave us the number of children eligible to attend each grade. As explained above, we assumed that each grade draws equally from the two most frequent ages in the grade, and only from those ages (for instance, 50% of second graders are 6 years old, and 50% are 7 years old). We also assumed that all demographic types for a given age grow at the same rate \mathcal{G}_{at} . Based on these assumptions, we were able to calculate the number of children of each *age* in each year as follows. Denote this number as N_{at} . Define $\mathcal{G}_{at} = N_{a,t} / N_{a,2000}$ for $t=2003, \dots, 2007$. We then calculated the household type measures $\mu_{l_{amt}}$ for $t=2003, \dots, 2007$ as $\mu_{l_{amt}} = \mu_{l_{am}} \mathcal{G}_{at}$, where $\mu_{l_{am}}$ are the Census 2000 measures whose calculation was described above.

School Choice Sets

The critical issue in the construction of choice sets is determining the public school associated to each block group for each grade and year. To arrive at this association, we first associated block groups with attendance zones, and then attendance zones with schools. We have a set of associations for 2003 and 2004 (prior to the attendance zone boundary changes), and another for 2005-2007. We refer to these as the 2003 and 2005 associations, respectively. The associations were done using GIS software.

While the task is relatively straightforward for 2005, it is not so for 2003. For instance, some attendance zones contain no schools, while others contain two; some attendance zones contain a school within their physical boundary that does not have the same name as the attendance zone; etc. To fix these problems, we defined the main school and the secondary school when needed for each attendance zone and school level. For a given attendance zone, the main school is the school whose name matches the attendance zone's, or the school that lies within the physical boundaries of the school and does not have the same name as the attendance zone, but does not have another attendance zone named after the school. The additional school (when it exists) is the school that does not have a specific attendance zone associated to it and lies within a zone that does have a main school. Overall, additional schools are quite rare in the data.

We determined main and additional schools based on conversations with GIS experts from DCPS. Those conversations also provided a solution to the problem that the attendance zones' map contains holes – i.e., some block groups are not associated to any attendance zone. We assigned those block groups to the closest public school for each level.

TABLE A1
Number of children by age

Age	Number
0	6306
1	6123
2	6352
3	6332
4	6657
5	6692
6	6985
7	7095
8	7068
9	7069
10	6887
11	6039
12	5674
13	5636
14	5426
15	5469
16	5578
17	5787
18	5595

Note: data from the 2000 Census. For each age, the count is an aggregate over Census-tract counts. For the count of 18-year olds, see Appendix IV.

TABLE A2
Number of Children by Age Bracket

	2000	2003	2004	2005	2006	2007	% Change 07 w.r.t 03
Total Population	571,744	577,777	579,796	582,049	583,978	586,409	1.49%
Children Below 5 yrs.	32,404	33,357	34,167	34,771	34,677	35,794	7.31%
Elem. Sch. Age Children (5-13)	59,746	59,561	57,186	55,078	53,873	52,068	-12.58%
High Sch.-Age Children (14-17)	22,504	23,145	24,534	25,156	26,049	26,198	13.19%

Note: for 2000, counts are from the 2000 Census. For the other years, counts are from Census Bureau estimates.

TABLE A3
Children Enrolled and Not Enrolled in School, by Age Bracket

	Enrolled	Not Enrolled	Total	Pct. Not Enrolled
Age 3-4	9,074	4,426	13,500	33.79%
Age 5-9	33,854	945	34,799	2.72%
Age 10-14	29,916	526	30,442	1.73%
Age 15-17	15,502	1,188	16,690	7.12%

Note: data from the 2000 Census, aggregated over Census tracts.

TABLE A4
Total Enrollment by Grade and Year

Grade	2003	2004	2005	2006	2007	% change 07 w.r.t. 03
PS	2,019	2,019	2,144	2,229	2,051	2%
PK	3,642	4,035	3,952	3,946	4,082	12%
K	6,630	6,392	6,558	6,294	5,908	-11%
1	6,923	6,735	6,417	6,606	6,293	-9%
2	6,806	6,419	6,354	6,096	6,121	-10%
3	6,656	6,400	6,032	5,944	5,781	-13%
4	7,002	6,387	6,133	5,780	5,658	-19%
5	7,021	6,647	6,063	5,890	5,567	-21%
6	6,770	6,564	6,343	5,822	5,686	-16%
7	6,504	6,465	6,259	6,079	5,608	-14%
8	5,988	6,075	6,048	5,905	5,795	-3%
9	6,655	7,070	6,882	7,146	6,896	4%
10	5,587	5,844	6,031	5,864	5,992	7%
11	4,711	4,915	4,998	5,079	5,000	6%
12	4,249	4,021	4,269	4,432	4,840	14%
TOTAL PS-12	87,160	85,983	84,480	83,110	81,277	-7%
Total G1-G8	53,670	51,692	49,649	48,122	46,509	-13%
Total G9-G12	21,202	21,850	22,180	22,521	22,728	7%
Census estimates - ages 5-13	59,561	57,186	55,078	53,873	52,068	-13%
Census estimates - ages 14-17	23,145	24,534	25,156	26,049	26,198	13%

Note: enrollment counts are based on our own calculations, aggregating over public, private and charter schools by grade and year. Rows with Census estimates for ages 5-13 and 14-17 are the same as in Table A2 and have been inserted here for comparison with enrollment totals for grades 1 through 8, and 9 through 12 respectively.