THE CURRENTS BENEATH THE "RISING TIDE" OF SCHOOL CHOICE: AN ANALYSIS OF STUDENT ENROLLMENT FLOWS IN THE CHICAGO PUBLIC SCHOOLS

M. Irmak Sirer,* Spiro Maroulis,* Roger Guimerà, Uri Wilensky, and Luís A. Nunes Amaral

Abstract

Existing research highlights that families face geographic, social, and psychological constraints that may limit the extent to which competition can take hold in school choice programs. In this paper, we address the implications of such findings by creating a network of student flows from 11 cohorts of eighth-grade students in the Chicago Public Schools (CPS). We applied a custom algorithm to group together schools with similar sending and receiving patterns, and calculated the difference in mean achievement between a student's attended and assigned high schools. For all identified school groupings, we found that the students were on average moving to higher achieving schools. We also found that the movement toward higher achieving schools of the top achievement quartile of students was over twice as large as that of the bottom quartile, but that the flows of both the highest and lowest achieving student quartiles were toward higher achieving destinations. Our results suggest that student movements in CPS between the years of 2001 to 2005 were consistent with creating market pressure for improvement as well as increasing segregation by achievement. However, further research into how schools responded to those movements is required to make inferences about the level or consequences of competition generated by choice-related reforms during that time.

INTRODUCTION

The idea that providing parents and students increased choice in schooling options will improve educational outcomes has been a leading theme in education reform (Berends et al., 2009; Gill et al., 2007). Proponents claim that increased choice provides the incentives necessary for schools to become more efficient in converting resources into outcomes (Chubb & Moe, 1990). From the perspective of supporters, choice will not only help those who exercise it, but also provide competition that will lead to system-wide improvements—the proverbial "rising tide that lifts all boats" (Hoxby, 2003). Opponents worry that school choice will not bring about the desired improvements but instead result in adverse distributional consequences, including increased academic and socioeconomic segregation (Fiske & Ladd, 2000). From the perspective of choice critics, already underserved students will be less likely or able to exercise choice, consequently exacerbating existing inequities in the system.

*These authors contributed equally to this work and are co-lead authors.

Journal of Policy Analysis and Management, Vol. 0, No. 0, 1–21 (2015)

Published by Wiley Periodicals, Inc. View this article online at wileyonlinelibrary.com/journal/pam
DOI:10.1002/pam.21826

A key aspect in analyzing the consequences of competition in choice programs is understanding how parents and students choose schools. Although a large literature exists on how parents and students choose schools, the implications of this research for studying competition have yet to be fully investigated. Two broad findings warrant particular consideration. First, the academic superiority of schools is just one of many factors considered by households when selecting a school (Armor & Peiser, 1998; Schneider et al., 1998). This is important because the less households value the academic characteristics of schools, the less likely it is that competition between schools on the basis of academic performance will emerge. To be clear, heterogeneity in preferences per se is not problematic. Increased choice may very well improve parental satisfaction and welfare by allowing families to sort themselves into schools that better match their preferences. However, to the extent that the policy expectation of a choice program is an across-the-board increase in academic achievement, low-achieving schools must observe students flowing out of their schools and toward higher achieving ones in order to feel pressure for improvement.

Second, households face geographic, social, and psychological constraints that can effectively limit the set of schools they consider (Bell, 2009; Lauen, 2007). This points to an often overlooked challenge in analyzing competition—defining the relevant "educational market" to use for analysis. The convention is to use metropolitan statistical areas (MSAs) as defined by the U.S. Office of Management and Budget or the geographic boundaries of school districts. However, MSAs and school districts are often large and heterogeneous. Given the known and unknown constraints facing households in a metropolitan area, precisely how "local" one should define an education market is not clear.

In this paper, we address these implications by conceptualizing a school district as a networked system of interacting parts (Maroulis et al., 2010), an approach akin to, but more general than, spatial models that explicitly model origin-to-destination migrations (Cooke & Boyle, 2011). Using data from 11 cohorts of eighth-grade students in the Chicago Public Schools (CPS) from 1994 to 2005, we create a network of between-school enrollment patterns for one of the largest public choice systems in the country. In contrast to current research, our focus is neither on investigating whether a subset of that network (choice schools) is better than another, nor on examining if a measure of the market structure that emerges from the interactions between elements of that network (e.g., the Herfindahl index) correlates with the performance of that system. Instead, we leverage the information contained in existing administrative data to ask a complementary question: To what extent are the student flows in this network consistent with the ones required to bring about market pressure for improvement?

To overcome the market boundary definition problem, we apply a module identification algorithm to the network data that identifies emergent subdistricts within CPS by grouping together schools with similar sending and receiving patterns in the network. Importantly, the algorithm makes no a priori assumptions about the factors that might limit student mobility. We identified 12 separate groups of schools within CPS with statistically equivalent flows, indicating that students flow in a manner that is much more "local" than the district boundaries. To characterize the extent to which students flow in a manner consistent with the prediction of economic theory (Tiebout, 1956), we define the school achievement differential (SAD) as the difference in mean achievement between a student's attended and assigned schools, and calculate and compare the mean SAD within and across groups of interest. We found that the mean SAD for all subdistricts was positive, indicating that students within each subdistrict were, on average, moving to higher achieving schools. When disaggregating SAD by prior student achievement, we found that the mean SAD of students in the highest achievement quartile was over twice as large as that of the lowest quartile, but that the flows of each group were on average toward schools with higher achievement. Our results suggest that the demand side of the educational market that has emerged from the open-enrollment program operates in a manner consistent with creating market pressure for improvement as well as increasing segregation of students by achievement. It is important to note, however, that our analysis focused only on the demand side of the system. Further research into the supply-side response to student movement is required to make inferences about the level or consequences of competition generated by choice-related reforms in CPS during that time.

BACKGROUND

Achievement Effects and School Choice

Theoretically, choice programs can lead to academic performance improvement through two broad mechanisms. One mechanism involves students sorting themselves into "better" schools. The schools can be better either in an absolute sense, such as having a higher value-added for all students; or in a relative sense, in that students select schools that are better matches for their learning styles or particular needs. Achievement gains through sorting do not necessarily require changes at existing schools in response to the program, only that higher quality schools that can accept students exist. Consequently, it is possible for the empirical effects of sorting to manifest themselves in the short term, particularly in studies that take advantage of the lotteries put in place to deal with oversubscription in pilot voucher programs, popular schools in public open-enrollment programs, and sought-after charter schools (Angrist et al., 2012; Bloom & Unterman, 2014; Cullen et al., 2006; Wolfe et al., 2013).

The results with respect to test score improvement in such studies have been mixed, which is perhaps not surprising given the broad range of schools that have been investigated. For example, analyses of voucher pilots in Milwaukee (Greene et al., 1997), New York City, Dayton, and Washington, DC (Howell et al., 2002) have found improvements in achievement for some subgroups of students, and lottery-based studies of charter schools have found substantial positive effects for some schools (Abdulkadiroglu et al., 2009; Angrist et al., 2012; Hoxby & Rockoff, 2005). However, analyses of broader samples of charter schools have resulted in more mixed findings (Bifulco & Ladd, 2006; Gleason et al., 2010; Sass, 2006; Zimmer et al., 2009), and an analysis using data collected from one of the largest public choice programs in the country finds that lottery winners at high schools in Chicago's open-enrollment program did not experience test score improvement (Cullen et al., 2006).

A second, and potentially farther reaching mechanism, involves a competitive process: when given a choice that is not dependent on residence, students will flow from low-performing schools to better ones. Schools losing students feel pressure to change in order to attract and keep students, which, in turn, creates pressure for all schools to change. In this way, the flow of students to better performing schools initiates a cycle of competition that can lead to system-wide improvement (Hoxby, 2003). How well this process works to increase achievement depends on both demand-side and supply-side factors. On the demand side, it depends on the extent to which students actually move to better schools. The more that students migrate toward better schools, the more likely it is that low-performing schools will feel pressure to improve. On the supply side, it depends on the extent to which school leaders, teachers, and other members of a school's community are willing or able to respond to these movements. Their response could be hindered either on account of inadequate incentives—such as funding scenarios where schools are not significantly penalized for the loss of students—or organizational and bureaucratic obstacles to change (Chubb & Moe, 1990).

Gains from competition are likely to take longer to unfold. Consequently, much empirical evidence on the competitive effects of choice comes from investigations attempting to connect metropolitan-level schooling productivity to the competitive process that arises when households choose local school districts via their residential decisions—a process that predates modern school choice programs and is referred to in the public finance literature as Tiebout choice (Oates, 2006; Tiebout, 1956). Economic theory predicts that as the level of Tiebout choice increases, so should a school's incentive to be more productive (i.e., produce more achievement per dollar of expenditure). Studies that test the impact of a Tiebout competition in the context of school choice typically operationalize competition through the creation of indices that measure the concentration of student enrollments or the availability of private schools. Variation across metropolitan areas in such measures is then correlated to system-wide performance. The findings from this work have been mixed (for review, see Belfield & Levin, 2002) and controversial (e.g., Hoxby, 2005; Rothstein, 2005), largely on account of the methodological difficulties associated with drawing causal inference from nonexperimental estimates (Bifulco, 2012). The primary issue is that concentration ratios, such as the Herfindahl index, that serve as indicators of competition in a metropolitan area, emerge as a result of student enrollment decisions that are likely not independent of the performance of the districts and schools within that area (Hoxby, 2000). Note, however, that even with more certainty in the causal effects from such studies, the issue of market boundary definition remains: an MSA or single school district may not provide the most appropriate boundaries for a population of households facing a variety of known and unknown constraints limiting which school they can choose.

Preferences and Constraints in Choosing Schools

In addition to work attempting to estimate the performance effects of the choice schools and programs, there is also a considerable body of empirical work investigating how households choose schools. Surveys investigating the preferences and attitudes that underlie choice behavior consistently find that parents value high academic achievement in schools, but also that a preference for academic superiority is just one of many factors considered by households (Armor & Peiser, 1998; Schneider et al., 1998). The relative importance of academic achievement in comparison to other factors is unclear, with studies finding that parents also highly value convenience, discipline, safety, and a match with their values (Armor & Peiser, 1998; Hastings et al., 2005; Lee et al., 1996). Additionally, there is evidence that parents do not always choose schools in accordance with their stated preferences. For instance, in an analysis of San Antonio charter school choosers, Weiher and Tedin (2002) show that over 60 percent of parents who ranked test scores as a top consideration moved their children into schools with worse scores.

Studies have also found a number of factors that potentially constrain choice decisions. While low-income families are generally in favor of choice programs (Lee et al., 1996; Plank et al., 1993), students from households with greater resources will be more likely to take advantage of them (Armor & Peiser, 1998; Carnegie Foundation, 1992; Martinez et al., 1996). This holds true even within populations of low-income students. Comparisons from voucher programs targeting students from low-income families in Milwaukee, Dayton, Washington, DC, New York City, and Cleveland show that the lower the education level of the mother of the household, the lower the likelihood that the family took advantage of the voucher program (see Gill et al., 2007, chapter 5, for review). Neighborhood-level disadvantage can constrain choice as well. Using administrative data similar to what we use in this study, Lauen (2007) finds that disadvantaged contexts reduce the likelihood that Chicago students opt out of attending their assigned high school for a private or (public) selective enrollment school. Lauen (2007) also finds that being assigned to a high-quality high school reduces the likelihood that a student will choose to enroll in a private, selective, or different non-neighborhood school. Finally, geographic considerations matter in multiple ways. Logistically, concerns such as transportation and commute times are often highlighted as primary constraints to mobility (Plank et al., 1993). Psychologically, households can develop preferences for neighborhoods and communities that frame their decisions (Bell, 2009).

Chicago Public Schools

As with other large urban school systems, CPS is no stranger to education reform. Major reform efforts date back to 1988 when the Illinois state legislature passed the Chicago School Reform Act. The act granted much greater planning and budget authority to local communities through the creation of local school councils (LSCs) composed of the school principal and teachers elected by school staff, as well as parents and community leaders elected by members of the community. While LSCs are still in place today, in 1995 the state legislature rewrote the act to grant authority over CPS to Chicago's mayor, who created a new CEO position. Paul Vallas, the first CEO from 1995 to 2001, led an effort of reforms based on testing and test performance. New graduation requirements, strict performance standards based on standardized test scores, and a required minimum score on the Iowa Tests of Basic Skills (ITBS) for high school enrollment were introduced. Schools with low scores were put on probation and often specifically targeted for reform. Arne Duncan, CEO from 2002 to 2009, augmented these reform efforts with a renewed focus on closing schools with poor performance and opening a number of new, mostly charter, schools.[1]

While the diversity of schools available to students and parents greatly expanded during this time period, school choice in CPS has its origins in efforts to desegregate Chicago's schools that predate recent reforms. In response to a 1980 consent decree signed with the U.S. Department of Justice, CPS began developing magnet programs within existing public schools, and new stand-alone magnet schools, to help alleviate the racial segregation. Public school choice in Chicago began to take on its current form in the 1994 to 1995 school year when open enrollment was introduced. Open enrollment gave parents and students the ability to attend any other neighborhood public school outside their geographically determined attendance area. Neighborhood schools must still give priority to students within their geographic attendance boundary, but students assigned to other neighborhood schools in the district can fill excess slots. A computerized lottery is used to handle cases of oversubscription at highly sought-after schools. This is the system that was in place for the cohorts of incoming high school students we analyze in this paper from 1994–1995 to 2004–2005, and is largely still in place today. Table 1 provides an overview of the quantity and variety of school types available to each cohort during this time, as well as the fraction of each cohort that elected to attend a school other than the attendance area school they were assigned.

DATA AND METHODS

The logic of school choice reform is predicated on the idea that student movements away from low-performing schools will serve as the impetus to competition among

[1] For more detailed history and analysis of reform in CPS, see Luppescu et al. (2011).

Table 1. CPS descriptive statistics, 1994 to 2005.

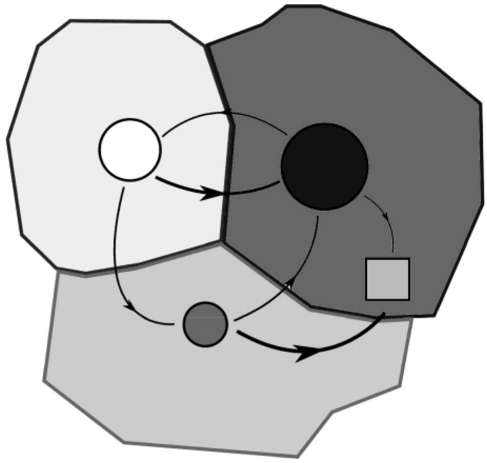
Cohort	Number of ninth graders	Total number of schools	Number of admissions criteria schools	Fraction leaving assigned school
1994 to 1995	24,071	62	22	0.514
1995 to 1996	22,784	61	21	0.519
1996 to 1997	22,282	64	23	0.501
1997 to 1998	20,215	65	23	0.539
1998 to 1999	14,200	68	23	0.567
1999 to 2000	16,013	70	25	0.539
2000 to 2001	15,249	72	26	0.573
2001 to 2002	15,808	73	26	0.562
2002 to 2003	15,767	76	28	0.535
2003 to 2004	15,773	83	31	0.551
2004 to 2005	17,924	91	30	0.564
All years	200,086	97	35	0.538

a substantial fraction of schools within reasonable proximity of each other. While students can benefit from choice in ways that do not involve competition (by sorting themselves into schools that are better "matches," for instance), the mechanism of competition is the linchpin of the argument the choice will lead to large-scale improvement. In this section, we describe how we apply network analytic techniques (Frank, 1995; Guimera & Amaral, 2005; Stouffer et al., 2012) to existing administrative data in order to better understand the extent to which student movements are consistent with the movements needed to bring about this mechanism.

Our analysis has three primary components. First, we use the CPS data to create a network of enrollment flows that enables us to define and visualize student mobility. Second, in order to tackle the challenge of defining the "market" boundaries in a large and heterogeneous district, we apply a module identification algorithm to the network data that groups together schools of similar sending and receiving patterns. Our algorithm provides us with a high-level map of subdistrict and student flows that we can use to organize, simplify, and disaggregate the large volume of student enrollment outcomes in the district. Third, we define and calculate a measure—SAD—to characterize and compare the extent to which students move from low- to high-performing schools across different parts of the district. After presenting a description of our data, we discuss each of these components in detail below.

CPS Data

The data cover 11 cohorts of students enrolling in ninth grade for the first time in a CPS high school for the academic years 1994–1995 to 2004–2005. This consists of 200,086 students in total. The primary sources of the data are student-level administrative and test score files from CPS. The administrative file contains the neighborhood high school to which students were assigned based on the location of their residence in the spring of their eighth-grade year, as well as the school they actually attended in the fall. We do not have information on the schools to which the student applied, only on the enrollment outcome for that student. With respect to student achievement, we have student performance on the ITBS in the eighth grade for all cohorts. For the cohorts from 2001 to 2005, we also have student-level performance on the Prairie State Achievement Examination (PSAE) scores in 11th grade. We do not have data on student performance in high school before 2001, which limits our SAD analysis to the years between 2001 and 2005. There are 97 high



Notes: Circles represent neighborhood schools with students assigned from their respective attendance areas. The square represents a school, such as a magnet or charter, that starts out without any assigned students. Arcs depict the net flows of students that were assigned to one school but enrolled in another. The thickness of an arc is proportional to the number of students moving from one school to the other.

Figure 1. Diagram of the Student Enrollment Network.

schools in the system during the time period we study. We placed those schools into one of two categories. Schools that had any programs that required an application requesting information from the students suggesting the possibility of selection on high (or low) academic performance were placed in the "admissions criteria" category. This category included, but was not limited to, the "selective enrollment" high schools in Chicago that accept only the top-performing students. Schools without selective programs were classified as "non-admissions criteria" schools.[2]

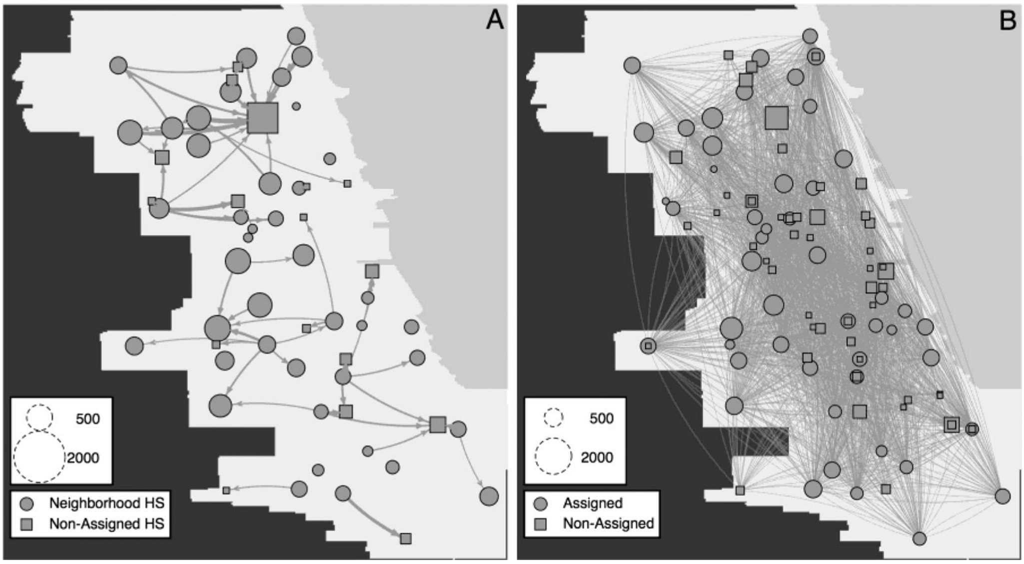
Constructing the Enrollment Network

The students' choices between schools for year y can be represented by a network of flows, as illustrated in Figure 1 for four schools. Each school i is connected to school j by an arc if there are students that were assigned to attend ninth grade in school i but instead enrolled in j. The number of such students defines the weight $w_{ij}(y)$ of the arc. We build this network for all the schools and students each academic year in our data, as well as an aggregate network across all years (Figure 2). In the aggregate network, the weights w_{ij} of the arcs represent the average yearly number of students who were assigned to i but enrolled in j for the years both schools were present.

Identifying Primary Flow Patterns and Emergent Subdistricts

In order to abstract from the complexity portrayed in Figure 2, we customized and applied a method of identifying groups, or "modules," of schools with similar

[2] A total of 21.1 percent of the student records were missing IOWA scores. A total of 18.6 percent of the student records between 2001 and 2005 were missing PSAE scores (i.e., the years for which PSAE scores were available). We did not observe any systematic patterns with respect to missing data across schools. Comparing the achievement of students who were only missing one of the two tests (e.g., comparing the IOWA scores of students missing PSAE vs. those not missing the PSAE, and vice versa), we observed that the mean achievement of students missing a test was slightly lower than the mean of the students not missing one, but neither difference was significant at the 5 percent level.



Notes: (A) School choice network for the Chicago Public School system, 2001. Only flows of greater than 40 students are shown. (B) Aggregate school choice network for the CPS system, 1994 to 2005.

Figure 2. School Choice Networks.

sending and receiving patterns. This allowed us to subsequently compare student flows within different modules, as well as investigate the stability of the student flows over time.[3] Constructing the modules involved a multistep process in which we (1) grouped schools with similar outflows, that is, schools with students leaving for a common set of other schools; (2) grouped schools with similar inflows, that is, schools receiving students from a common set of other schools; and then (3) intersected the outflow- and inflow-defined groups to identify modules of schools with statistically equivalent sending and receiving patterns. Seeking these modules required defining criteria for evaluating different partitions of schools, and then using the probabilistic optimization procedure of simulated annealing to identify the optimal grouping (Guimera et al., 2007; Stouffer et al., 2012).

More precisely, in a network with n schools, the outflows of a school i can be described by the vector $\vec{w}_i^{out} = \{w_{i1}, w_{i2}, \ldots, w_{in}\}$. In order to find groups of schools with similar outflows, we consider the pairwise similarity $\vec{w}_i^{out} \cdot \vec{w}_j^{out}$ between all pairs of schools. Any given grouping, P, of schools into subdistricts can be assigned a score, $\Pi^{\text{out}}(P)$, defined as follows:

$$\Pi^{out}(P) = \sum_{groups} ([\textit{within group similarity}] - [\textit{within group similarity expected from chance}]). \quad (1)$$

[3] We compared the similarity of the network partitions in each of the years to each other using an information entropy based metric, normalized mutual information (NMI). The flow patterns are stable over time, which allows us to investigate the average of these yearly networks. NMI is established as the standard tool for comparing the similarity of two network partitions. For an in-depth description, see Danon et al. (2005).

Using simulated annealing, we identified the grouping of schools that maximized Π^{out} (Guimera & Amaral, 2005; Guimera et al., 2007). This resulted in groups of schools that were similar in the sense they lost students to the same other schools. We repeated this process using the similarity of inflows, $\vec{w}_i^{in} = \{w_{i1}, w_{i2}, \ldots, w_{in}\}$ to find an analogous grouping of schools, where the groups were defined by gaining students from the same sources. We then took the intersection of the inflow and outflow groups to find modules of schools that are equivalent in both. One can think of these modules as the naturally emerging subdistricts of schools that have students flowing both within and between them. Importantly, no a priori assumptions about those boundaries were used. Statistical similarities in the data alone determine the groups (see Appendix for additional details).[4]

Characterizing Student Flows

To characterize the extent to which students move from low- to high-performing schools in this network, we define a new measure—SAD. For each student, SAD is defined as the difference between the mean PSAE of their attended school j and their assigned school i. We then calculated the mean SAD:

- over all students in the district who did not enroll in their assigned schools in the 2001–2002 to 2004–2005 cohorts (the ones for which we had PSAE scores);
- disaggregated by emergent "subdistricts" within CPS, as defined above;
- disaggregated by intra- and intersubdistrict flows;
- disaggregated by the eighth-grade ITBS score of the students.

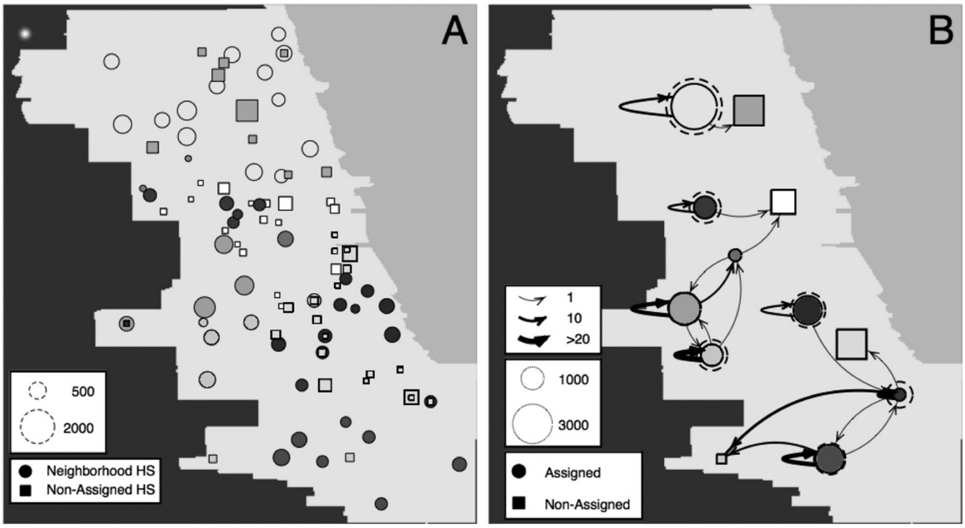
When aggregating SAD over an entire district or subdistrict, one can interpret the mean SAD as providing an indicator of demand-side pressure for improvement for that (sub)district. A positive mean SAD is consistent with the movement required to generate market pressure and spur competition; a zero or negative mean SAD indicates the opposite. To test if the mean SAD values were significantly different from the values expected from chance alone, we constructed a null model where students were initially associated with their assigned school, and then randomly chose a school to attend. The probability that any given student selected a school was proportional to the school's size. We then simulated the null model 100,00 times to construct a reference distribution of outcomes against which we could compare our observed outcomes.

RESULTS

Subdistrict Identification

Applying our module identification algorithm to the aggregate CPS enrollment network (Figure 2), we identified 12 subdistricts of statistically similar sending and receiving patterns within CPS. Four of the 12 subdistricts are comprised entirely of schools, such as charter schools, that were not assigned any students so their statistical similarity is based entirely on receiving students that were assigned to similar schools. The subdistricts identified by our algorithm and the main flow patterns among them are shown in Figure 3. One can think of Figure 3 as a high-level "map" of CPS that will help us decompose and interpret the district-level SAD results.

[4] All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's Web site and use the search engine to locate the article at http://www3.interscience.wiley.com/cgibin/jhome/34787.



Notes: (A) Subdistrict organization of CPS schools. Two schools shown with the same shade have similar student sending and receiving patterns. (B) Map of flow biases previously hidden in the aggregate school choice network (Figure 2B). Each emergent subdistrict is shown as a node positioned close to its center of mass. Node size shows yearly average of total enrolled students. Segmented line around a node shows yearly average of assigned students. Arcs represent student flows that are at least one standard deviation more likely than the null model. Arc thickness corresponds to the number of standard deviations the flow differs from the null model.

Figure 3. Flow Patterns in the Aggregate Network.

In panel A of Figure 3, schools with the same shade belong to the same emergent subdistrict. The shape of the school indicates whether the school was assigned any students or not. Circles represent schools that were assigned students; while squares represent schools that were not assigned any students. The size of the school is proportional to its enrollment. Individual schools in panel A are placed at their geographic location. In panel B, we replace the individual schools with a single node representing their emergent subdistrict corresponding in shade to the subdistricts in panel A. The node for any given subdistrict is placed in the center mass of the schools that comprise it. Square nodes represent the four subdistricts comprised entirely of schools that were not assigned students. The arrows, or arcs, highlight the "highways" of student flows—that is, the direction in which students are more likely to travel in comparison to random expectation. The arc thickness is proportional to the number of standard deviations by which the flows differ from random expectation. For clarity, only flow patterns with deviations higher than one standard deviation are shown.

The first thing to notice about Figure 3 is that the subdistricts in panel B are largely isolated from each other. The biggest deviations from random expectation—the thickest arcs—are the self-arcs that feed back into the same subdistrict from which they originated. These self-arcs represent schools in the same subdistrict trading students among each other. Panel A suggests that this deviation is at least in part related to spatial considerations—schools of the same shade tend to be geographically close to each other. This is not surprising. Interestingly, however, there are also several instances of schools that are very close to each other but belonging to different subdistricts (i.e., a different shade), suggesting that not all the clustering is due to geography. Regardless of the underlying reasons, the strong deviation from the null model indicates that students flow in a manner that is much more "local" than the district boundaries.

Table 2. Mean SAD of intra- and intersubdistrict student flows.

	All students	Students attending nonadmissions criteria schools only
All flows	15.21 (16.7)	4.72 (6.2)
Between-subdistrict flows	18.77 (17.4)	5.92 (6.6)
Within-subdistrict flows	5.75 (9.6)	2.77 (4.9)

Notes: All SAD means are significantly different from the null model at the 0.01 level. Standard deviations in parentheses.

The second thing to notice about Figure 3 is that seven of the eight subdistricts containing assigned schools (the circles in panel B) are losing students, with the recipients often being subdistricts that are comprised entirely of nonassigned schools (the squares). This is illustrated in panel B through the segmented lines around the nodes. The segmented lines depict the yearly average of the number of students assigned to schools within the subdistrict. The size of the node is proportional to the yearly average of total students actually enrolled in a school in the subdistrict. Large differences between the segmented line and the full size of the node indicates when the subdistrict is a net exporter of students.

To summarize, the two most prominent deviations from the null model that emerge from the creation of our flow map are the self-arcs originating and ending in the same subdistrict, and the flows from subdistricts of mostly assigned students to nearby subdistricts entirely composed of schools with no assigned students. This strongly indicates that students flow in a much more circumscribed manner than one would expect if the most salient boundaries were the geographic district boundaries. However, it does not speak to the issue of whether these flows facilitate Tiebout-style pressure for improvement. To address this issue, we turn to the SAD calculations on the student enrollment network. First, we present the aggregate, district-level SAD results. Second, we interpret those results through the lens of the enrollment flow map presented in this section (Figure 3). In particular, we disaggregate the district-level SAD results by type of flow (self-arcs vs. between-subdistrict arcs) and subdistrict (Table 2). Third, we disaggregate the district-level SAD results by the eighth-grade achievement level of the students (Table 4).

District-Level Flows

Examining SAD, the difference between the mean PSAE score of students' attended school and assigned schools, at the district level reveals that on average student flows are consistent with the idea that choice can create market pressure for improvement. Since some migration is due to students who leave their assigned school to attend schools that select on the basis of prior achievement, we calculate SAD two ways—one using only students who attend nonadmissions criteria schools, and one using all students in the district. The mean SAD when averaging over only students who attend nonadmissions criteria schools is 4.72 points and statistically different from the expectation of the null model ($P < 0.0001$). This indicates that on average students who opt out of their assigned schools do indeed attend higher achieving schools. The mean SAD when including students who attend admissions criteria schools is 15.21 points. To put the magnitude of these numbers in perspective, 4.72 points translates to approximately 0.19 standard deviations on the distribution of PSAE scores from all students in our sample; and 15.21 points translates to 0.61 standard deviations.

Table 3. Student flows by subdistrict.

Emergent subdistrict	Number of schools	Number of assigned students	Fraction moving within subdistrict	Fraction leaving subdistrict	Mean SAD
1	6	1,806	0.161	0.529	15.01
2	9	2,292	0.12	0.489	12.62
3	2	1,232	0	0.860	15.2
4	7	2,043	0.233	0.309	9.42
5	15	5,864	0.181	0.362	21.48
6	5	2,215	0.102	0.256	14.1
7	4	2,154	0.084	0.504	11.02
8	1	318	0	0.214	22.43

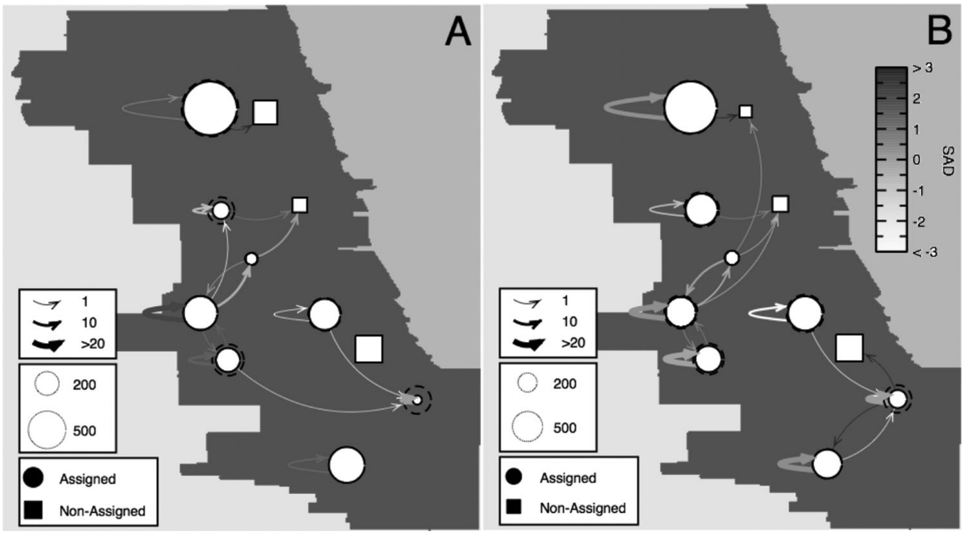
Notes: Subdistrict values were aggregated over all cohorts between 2001 and 2005 for students assigned to a school in the subdistrict. All SAD means are significantly different from the null model at the 0.01 level.

Intra- and Intersubdistrict Flows

Disaggregating the district-level SAD results by flow type, we see that students are finding higher achieving schools both within and outside of their subdistrict. The mean SAD for between-subdistrict flows is 18.77 points and for within-subdistrict flows is 5.75 points (Table 2). Recalculating both values after excluding students who attend admission-criteria schools, the values expectedly decline but remain statistically significant from the null model and are positive. The mean SAD using students from nonadmissions criteria schools is 5.92 points for between-subdistrict flows and 2.77 points for within-subdistrict flows.

Disaggregating the district-level SAD results by subdistrict paints a more nuanced picture. Table 3 reports the results. It gives the number of schools, number of students assigned to schools in the subdistrict, fraction of those students choosing another school within the same subdistrict as their assigned school, fraction of students choosing a school outside the subdistrict of their assigned school, and mean SAD of the students assigned to schools within the subdistrict but choosing not to attend their assigned school. Subdistrict values were aggregated over the 2001 to 2005 cohorts, as those were the ones for which PSAE scores from the high schools were available. Since SAD cannot be calculated for schools that do not have assigned students, Table 3 does not contain rows for the four subdistricts composed of entirely those types of schools.

Using the information in Table 3 to characterize the demand-side pressure for improvement faced by schools in each subdistrict, we see that the mean SAD of every subdistrict is positive and statistically different from the null model. This is consistent with the idea of creating pressure for improvement. However, disaggregating by subdistrict also reveals two important differences. First, we see that the subdistricts vary with respect to how much students move out of assigned schools. The fraction of students not attending their assigned school ranges from 0.214 to 0.86. Interestingly, the two most extreme values of that range belong to subdistricts (3 and 8) that contain a very small number of schools. Subdistrict 3 consists of two schools that lose all their moving students to higher performing schools outside the subdistrict (mean SAD, 15.2). Subdistrict 8 consists of only one school that also loses students to other higher performing schools (mean SAD, 22.43). But the fact that the school in subdistrict 8 keeps over 78 percent of its assigned students and that there are no other schools in the district with sending and receiving patterns that were statistically equivalent to it implies the existence of either unique attraction or constraint. Second, we see that not only do subdistricts vary with respect to which students opt



Notes: Nodes represent emergent subdistricts. Arc thickness denotes flow volume in terms of deviation from null model. Arc shade denotes mean SAD along an arc in terms of deviation from null model.

Figure 4. Flow Patterns for (A) High-Achieving Students and (B) Low-Achieving Students.

out of their assigned schools, but also vary in the extent to which students who opt out migrate toward schools with higher performance. The mean SADs range from 9.42 points to 22.43 and differ from each other much more than what one would expect from chance ($P < 0.001$ for a one-way ANOVA testing the equivalence of the means).

High Achiever versus Low Achiever Flows

We used the eighth-grade scores on the ITBS to create two groups of students. We denoted students in the top 25 percent of test scores as "high achieving students," and students in the bottom 35 percent of test scores as "low achieving students." The percentiles were set so that the volume of flow (total number of moving students) was equal between the two groups. The larger window for low-achieving students stemmed from the fact that low-achieving students did not flow as much. A much larger portion of the bottom quartile stayed at the initial school to which they were assigned. Figure 4 keeps the same subdistrict groupings from Figure 3 but separates the flow patterns of high- and low-achieving students. To adjust for the fact that schools with selective enrollment criteria were only available to high-achieving students, in Figure 4 we only included students who enrolled in nonadmissions criteria schools.[5] The subdistricts are again represented by circles, with the exception of subdistricts containing only schools with no assigned students, which are represented by squares. The arc between any two subdistricts denotes flow volume of the students traveling from the origin to destination subdistrict. Self-arcs that originate and end in the same circle denote flows of students who did not attend their assigned school but chose another school in the same subdistrict. The thicker the arc, the more the flow between origin and destination flow deviates from

[5] Note there is one less subdistrict in Figure 4 than in Figure 3. This is because one subdistrict consists of only selective schools.

Table 4. Mean school achievement differentials and intersubdistrict movers by achievement level.

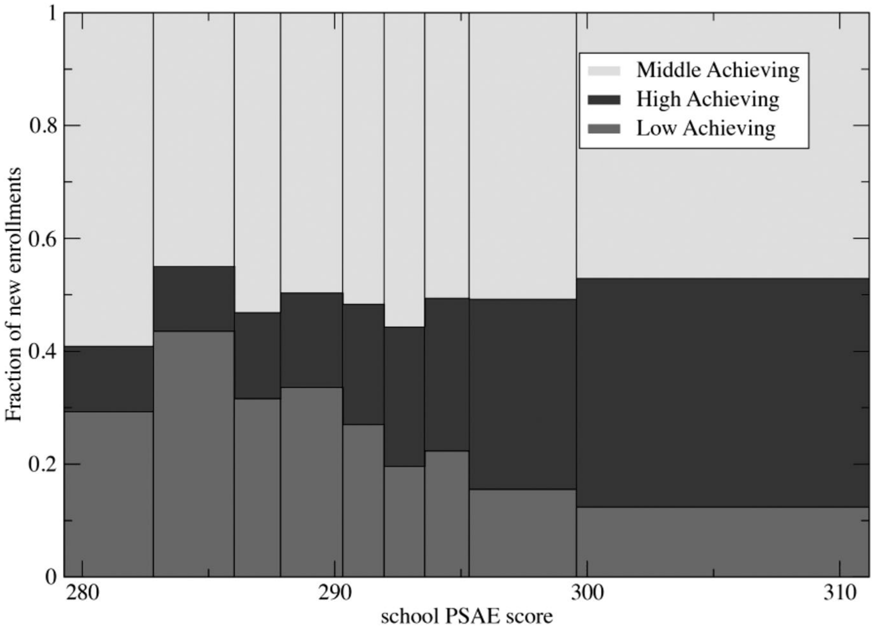
	High achieving	Low achieving	Difference	*P*-value
	All students			
Mean SAD (all)	27.7 (17.2)	4.99 (9.2)	22.71	<0.0001
Mean SAD (intersubdistrict arcs)	30.0 (17.1)	6.46 (9.1)	23.64	<0.0001
Frac. switching subdistricts	0.842	0.689	0.153	<0.0001
	Students attending nonadmissions criteria schools only			
Mean SAD (all)	6.07 (6.2)	2.72 (6.1)	3.35	<0.0001
Mean SAD (intersubdistrict arcs)	7.16 (6.4)	4.30 (6.6)	2.86	<0.0001
Frac. switching subdistricts	0.625	0.584	0.041	<0.0001

Note: Standard deviations in parentheses.

the expectation of the null model. For clarity of the visualization, only deviations of greater than one standard deviation are depicted. Arc shade denotes mean SAD of the students traveling from one subdistrict to another. As with the arc thickness, the arc shade is scaled with respect to its deviation from the null model.

Comparing the flow patterns in Figure 4 suggests several differences between the two groups. In particular, the low-achieving students appear to have thicker self-arcs with a lower mean SAD and intersubdistrict arcs with a greater number of destinations. We further investigated the differences implied by the flow map by quantifying the low and high achiever flows in terms of the SAD and the fraction of students leaving their subdistrict. Table 4 summarizes the results. The most important thing to note in Table 4 is that for both low and high achievers the mean SAD is positive, but the mean SAD for high achievers (6.02) is higher than that for low-achieving students (2.72). That is, both groups are flowing from lower to higher achieving schools, but high-achieving students do so to a greater extent. Additionally, low-achieving "movers" that opt out of their assigned school are less likely to choose a school outside their original subdistrict. The "Frac. Switching Subdistricts" rows in Table 4 indicate that 62.5 percent of the high-achieving students who move to nonadmissions criteria schools pick a school outside their subdistrict, as opposed to 58.4 percent of low-achieving students. The difference grows when students attending admissions criteria schools are included (84.2 percent for high-achieving movers vs. 68.9 percent for the low-achieving movers). In all cases the differences are significantly different at the 0.01 level.

To better understand the implication of the difference in flows between low- and high-achieving students, we examined the concentration of achievement both within schools and within subdistricts that occurs as a result. Using the same achievement percentiles to define groups as before, we characterized each nonadmissions criteria school by the fraction of low-, middle-, and high-achieving students that enrolled in that school in ninth grade in each year, and then compared this composition with the school's average PSAE score from the year previous to enrollment. We ranked all the schools according to their average PSAE score and binned them into groups of six schools. We then calculated the average low, medium, and high fractions of incoming ninth graders and the average school PSAE scores for each bin. (Students who enrolled in their assigned school are not included.) Figure 5 shows the composition of low-, middle-, and high-achieving students among incoming ninth graders for each bin of six schools. In accordance with the mean SAD results in Table 4, the fractions of incoming students in Figure 5 show that low-achieving students enrolled in schools with lower PSAE scores, compared to high-achieving



Notes: Fractions of low- and high-achieving students who enroll in schools with nonselective admissions criteria. Students who remain in their assigned school are not included. The *y*-axis shows the composition for each school at each year, while the *x*-axis shows the average PSAE score of the school from the previous year. Schools are ordered on the *x*-axis from lowest to highest PSAE score and each *x*-axis bin contains six schools. The three shaded areas in each bin refer to the fractions of low-, middle-, and high-achieving students for each bin.

Figure 5. Concentration of Student-Level Achievement within Schools.

students. This resulted in a concentration of low-achieving students in schools with lower scores, and a concentration of high-achieving students in schools with higher scores, even though none of the schools considered in this analysis selected students on the basis of achievement.

CONCLUDING COMMENTS

School choice is one of the most debated reform ideas in education. Underpinning the arguments of both proponents and critics are assumptions about changes in enrollment patterns brought about by school choice. We contribute to this debate by creating and analyzing the network of student flows from assigned to attending schools in a large urban district. In particular, we provide evidence that the demand side of the educational market that has emerged from the open-enrollment program in Chicago is indeed operating in a manner consistent with creating Tiebout-style pressure for improvement. Students are on average migrating to higher achieving schools, even when including only students that attended nonselective schools in the analysis. While this pattern of student movement is predicted by economic theory, to the best of our knowledge no one has ever directly quantified and investigated these flows in the context of a public choice system. Disparities in the likelihood to opt out of assigned schools have been reported (e.g., Lauen, 2007), as has the aggregate association between MSA-level measures of competition and performance (e.g., Hoxby, 2005). But a direct investigation of the underlying flows theoretically responsible for connecting those two pieces of the school choice puzzle has not been previously explored.

Moreover, we find differences in flows between high- and low-achieving students. This is consistent with prior work on CPS reporting that students in disadvantaged contexts were less likely to exercise school choice (Lauen, 2007). We extend that work by showing that conditional on choosing, higher initial achievement is also associated with selecting higher achieving destinations. Such a dyadic calculation is important because it helps us interpret the low achiever/high achiever difference in the context of understanding competition within school choice programs. That is, we are able to see that both high- and low-achieving students migrate in a manner consistent with creating Tiebout-style pressure for improvement, but that high-achieving students do so much more than low-achieving students (even after accounting for the fact that many high achievers attend selective schools).

In theory, there are two potential sources for this difference. One is that low-achieving students are less willing or able to find higher achieving schools. Another is that geographic proximity is an important factor in choosing a school, independent of a student's willingness or ability to identify high-performing schools. To the extent that there is initial spatial concentration in achievement, even an equal valuation of geographic proximity by low- and high-achieving students could yield a disparity in flows. Future work could attempt to parse the relative impact of these sources of difference by estimating a spatial interaction model (Cooke & Boyle, 2011) or network selection model (Frank, 2009; Handcock et al., 2008; Snijders, 2005) that predicts the flows between schools as a function of distance, SAD, and other factors. Our analysis demonstrates that—regardless of the underlying reasons—student-level enrollment choices have aggregated in a way where groups of schools "competing" for the same students (i.e., our subdistricts) may face different levels of demand-side pressure for improvement.

Our findings have implications for both researchers and policymakers. For researchers, our results highlight the importance of considering more "local" educational markets when analyzing competition in school choice. For example, with multicity data one could imagine investigating the association between common measures of competition, such as the Herfindahl index, and school productivity at the subdistrict level, as opposed to the MSA or district level. (We did not do so here on account of the small number of subdistricts.) For policymakers in districts with choice programs, conducting analyses such as ours can help identify sets of students and schools where student flows are least consistent with the aims of the program. For example, in our analysis, subdistrict 6 had the second-lowest fraction of students moving to another school and a SAD lower than many of the other subdistricts (Table 3). This would make it a likely target for further investigation or assistance for programs where a primary goal is promoting competition. Moreover, identifying such subdistricts can be done quickly and at low cost, as it only requires analyzing existing administrative data.

In drawing policy implications from our findings, it is crucial to note several limitations. First, the focus of our analysis was on characterizing a demand-side precursor to competition—directed student flows. We do not claim to know how schools responded to losses or gains in enrollment. If administrators in schools with net outflows responded by improving teacher quality, increasing academic standards, or implementing administrative changes to better serve the remaining students, then both our aggregated and conditional SAD results could be meaningfully interpreted as indicators of competition. If, on the other hand, political or organizational obstacles prevented meaningful change in those schools, our SAD measure is only useful in characterizing student demand and not the overall level of competition in the district. Second, while the open-enrollment program in Chicago certainly facilitated the movement of students between schools, it should not be interpreted as their singular cause. A large number of reform initiatives were undertaken in CPS between the academic years of 1994–1995 and 2004–2005. The enrollment patterns analyzed in this study were to some extent a consequence of all these efforts. Third, our analysis is limited to intradistrict choice among public schools. We do not have data on flows to private schools or to other districts in the Chicago metropolitan area (via residential relocation). However, we note that even if we did have those data, it is not clear, a priori, where we would want to draw the boundary on relevant "education market." Interdistrict, interschool, and intersector (public and private) competition are all important elements in understanding the dynamics of school choice (Neal, 2002), and as underscored in this study, households face contextual factors that can effectively limit the set of schools they consider (Lauen, 2007).

This final limitation points to an important methodological contribution of our paper. We have provided an approach for empirically identifying salient education markets that does not require imposing any assumptions about school sectors or constraints on student decisionmaking. By applying our module identification algorithm and general network approach to existing administrative data in many other districts and metropolitan areas, researchers and policymakers can gain a more nuanced understanding of the connection between student choices and market forces in education.

M. IRMAK SIRER is a Data Scientist at Datascope Analytics, 180 W Adams, Suite 801, Chicago, IL 60603 (e-mail: irmak.sirer@datascopeanalytics.com).

SPIRO MAROULIS is an Assistant Professor in the School of Public Affairs at Arizona State University, 411 North Central Avenue, Suite 400, Phoenix, AZ 85004 (e-mail: Spiro.Maroulis@asu.edu).

ROGER GUIMERÀ is a Research Professor at the Institució Catalana de Recerca i Estudis Avançats (ICREA), in Barcelona 08010, and the Departament d'Enginyeria Química at Universitat Rovira i Virgili, Tarragona 43007, Catalonia. He is also a member of the External Faculty at the Northwestern Institute on Complex Systems, Evanston, IL 60208 (e-mail: roger.guimera@urv.cat).

URI WILENSKY is a Professor of Learning Sciences, Computer Science, and Complex Systems at Northwestern University. He is a co-founder of the Northwestern Institute on Complex Systems and is the director of the Center for Connected Learning and Computer-Based Modeling, Annenberg Hall 337, 2120 Campus Drive, Evanston, IL 60208 (e-mail: uri@northwestern.edu).

LUÍS A. NUNES AMARAL is a Professor of Chemical and Biological Engineering, Medicine, and Physics and Astronomy, and co-director of the Northwestern Institute on Complex Systems at Northwestern University, 2145 Sheridan Road, Room E136, Evanston, IL 60208 (e-mail: amaral@northwestern.edu).

ACKNOWLEDGMENTS

The authors acknowledge the support of the National Science Foundation (HSD-0624318). We are also grateful to the Chicago Public Schools and the Consortium on Chicago School Research for their data assistance.

REFERENCES

Abdulkadiroglu, A., Cohodes, S., Dynarski, S., Fullerton, J., Kane, T. J., & Pathak, P. (2009). Informing the debate: Comparing Boston's charter, pilot and traditional schools (Tech. Rep.). Boston, MA: The Boston Foundation.

Angrist, J. D., Dynarski, S., Kane, T. J., Pathak, P., & Walters, C. R. (2012). Who benefits from KIPP? Journal of Policy Analysis and Management, 31, 837–860.

Armor, D., & Peiser, B. (1998). Interdistrict choice in Massachusetts. In P. Peterson & B. C. Hassel (Eds.), Learning from school choice (pp. 157–186). Washington, DC: Brookings Institution Press.

Belfield, C. R., & Levin, H. M. (2002). The effects of competition between schools on educational outcomes: A review for the United States. Review of Educational Research, 72, 279–341.

Bell, C. (2009). Geography in parental choice. American Journal of Education, 115, 493–521.

Berends, M., Springer, M. G., Ballou, D., & Walberg, H. (Eds.). (2009). Handbook of research on school choice. London: Routledge, Taylor and Francis Group.

Bifulco, R. (2012). Can nonexperimental estimates replicate estimates based on random assignment in evaluations of school choice? A within-study comparison. Journal of Policy Analysis and Management, 31, 729–751.

Bifulco, R., & Ladd, H. F. (2006). The impact of charter schools on student achievement. Education Finance and Policy, 1, 50–90.

Bloom, H. S., & Unterman, R. (2014). Can small high schools of choice improve educational prospects for disadvantages students? Journal of Policy Analysis and Management, 33, 290–319.

Carnegie Foundation. (1992). School choice: A special report (Tech. Rep.). Stanford, CA: The Carnegie Foundation for the Advancement of Teaching.

Chubb, J., & Moe, T. (1990). Politics, markets, and America's schools. Washington, DC: The Brookings Institution Press.

Cooke, T. J., & Boyle, P. (2011). The migration of high school graduates to college. Educational Evaluation and Policy Analysis, 33, 202–213.

Cullen, J. B., Jacob, B. A., & Levitt, S. (2006). The effect of school choice on participants: Evidence from randomized lotteries. Econometrica, 74, 1191–1230.

Danon, L., Diaz-Guilera, A., Duch, J., & Arenas, A. (2005). Comparing community structure identification. Journal of Statistical Mechanics: Theory and Experiment, 2005, P09008.

Fiske, E. B., & Ladd, H. F. (2000). When schools compete: A cautionary tale. Washington, DC: Brookings Institution.

Frank, K. A. (1995). Identifying cohesive subgroups. Social Networks, 17, 27–56.

Frank, K. A. (2009). Quasi-ties directing resources to members of a collective. American Behavioral Scientist, 55, 1613–1645.

Gill, B., Timpane, P. M., Ross, K., Brewer, D. J., & Booker, K. (2007). Rhetoric versus reality: What we know and what we need to know about vouchers and charter schools. Santa Monica, CA: RAND Corporation.

Gleason, P., Clark, M., Tuttle, C. C., & Dwoyer, E. (2010). The Evaluation of Charter School Impacts: Final Report. NCEE 2010–4029. Washington, DC: National Center for Education Evaluation and Regional Assistance.

Greene, J. P., Peterson, P. E., & Du, J. (1997). The effectiveness of school choice: The Milwaukee experiment. Cambridge, MA: Program in Education Policy and Governance and Center for American Political Studies, Harvard University.

Guimera, R., & Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. Science, 433, 895–900.

Guimera, R., Sales-Pardo, M., & Amaral, L. A. N. (2007). Module identification in bipartite and directed networks. Physical Review E, 76, 036102.

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). Statnet: Software tools for the representation, visualization, analysis and simulation of network data. Journal of Statistical Software, 24, 1548.

Hastings, J. S., Kane, T. J., & Staiger, D. O. (2005). Parental preferences and school competition: Evidence from a public school choice program (No. w11805). Cambridge, MA: National Bureau of Economic Research.

Howell, W., Wolfe, P., Campbell, D., & Peterson, P. (2002). School vouchers and academic performance: Results from three randomized field trials. Journal of Policy Analysis and Management, 21, 191–217.

Hoxby, C. M. (2000). Does competition among public schools benefit students and taxpayers. American Economic Review, 90, 1209–1238.

Hoxby, C. M. (2003). School choice and school productivity: Could school choice be a tide that lifts all boats? In C. M. Hoxby (Ed.), The economics of school choice. Chicago, IL: The University of Chicago Press.

Hoxby, C. M. (2005). Competition among public schools: A reply to Rothstein (2004), No. 11216. Cambridge, MA: National Bureau of Economic Research.

Hoxby, C. M., & Rockoff, J. E. (2005). Findings from the city of big shoulders. Education Next, 4, 52–59.

Lauen, D. L. (2007). Contextual explanations of school choice. Sociology of Education, 80, 179–209.

Lee, V. E., Croninger, R. G., & Smith, J. B. (1996). Equity and choice in Detroit. In B. Fuller & R. Elmore (Eds.), Who chooses? Who loses? Culture, institutions, and the unequal effects of school choice. New York, NY: Teachers College Press.

Luppescu, S., Allensworth, E. M., Moore, P., dela Torre, M., Murphy, J., & Jagesic, S. (2011). Trends in Chicago's schools across three eras of reform (Tech. Rep.). Chicago, IL: University of Chicago, The Consortium on Chicago School Research.

Maroulis, S., Guimera, R., Petry, H., Stringer, M., Amaral, L., Gomez, L., & Wilensky, U. (2010). Complex systems view of educational policy research. Science, 330, 38–39.

Martinez, V., Godwin, K., & Kemerer, F. (1996). Public school choice in San Antonio: Who chooses and with what effects? In B. Fuller & R. Elmore (Eds.), Who chooses? Who loses? Culture, institutions, and the unequal effects of school choice. New York, NY: Teachers College Press.

Neal, D. (2002). How vouchers could change the market for education. Journal of Economic Perspectives, 16, 25–44.

Oates, W. E. (2006). The many faces of the Tiebout model. In W. A. Fischel (Ed.), The Tiebout model at fifty (pp. 21–45). Cambridge, MA: Lincoln Institute of Land Policy.

Plank, S., Schiller, K., Schneider, B., & Coleman, J. S. (1993). Effect of choice in education. In E. Rasell & R. Rothstein (Eds.), School choice: Examining the evidence. Washington, DC: Economic Policy Institute.

Rothstein, J. (2005). Does competition among public schools benefit students and taxpayers? A comment on Hoxby (2000), No. 11215. Cambridge, MA: National Bureau of Economic Research.

Sass, T. R. (2006). Charter schools and student achievement in Florida. Education Finance and Policy, 1, 91–122.

Schneider, M., Teske, P., Marschall, M., & Roch, C. (1998). Shopping for schools: In the land of the blind, the one-eyed parent may be enough. American Journal of Political Science, 42, 769–793.

Snijders, T. A. (2005). Models for longitudinal network data. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), Models and methods in social network analysis (pp. 215–247). Cambridge, MA; New York, NY: Cambridge University Press.

Stouffer, D., Sales-Pardo, M., Sirer, M. I., & Bascompte, J. (2012). Evolutionary conservation of species' roles in food webs. Science, 335, 1489–1492.

Tiebout, C. (1956). A pure theory of local expenditures. Journal of Political Economy, 64, 416–424.

Weiher, G. R., & Tedin, K. L. (2002). Does choice lead to racially distinctive schools? Charter schools and household preferences. Journal of Policy Analysis and Management, 21(1), 79–92

Wolfe, P., Kisida, B., Gutmann, B., Puma, M., & Rizzo, L. (2013). School vouchers and student outcomes: Experimental evidence from Washington, DC. Journal of Policy Analysis and Management, 32, 246–270

Zimmer, R. W., Gill, B., Booker, K., Lavertu, S., Sass, T. R., & Witte, J. F. (2009). Charter schools in eight states: Effects on achievement, attainment, integration, and competition. Santa Monica, CA: RAND Corporation.

APPENDIX

Identifying Modules

We define the outflow profile of a school "i" as the vector $\vec{w}_i^{out} = \{w_{i1}, w_{i2}, \ldots, w_{in}\}$. Following Stouffer et al. (2012), the pairwise similarity between schools i and j is given by $\vec{w}_i^{out}.\vec{w}_j^{out}$, and the similarity of any given group (module) is calculated by summing the pairwise similarity between all pairs i and j of schools in the group. The intuition underlying our approach is to find the groups of schools that maximize within-group similarity. However, note that schools with many students leaving will have higher values in their w_i and therefore are likely to share more outgoing students in the same direction with others. Similarly, target schools in which many students enroll introduce a bias toward coinciding targets. To account for this bias, rather than simply maximizing within-group similarity, we look for the partitioning of schools with within-group similarity that maximally deviates from the similarity expected by chance (Guimera & Amaral, 2005; Guimera et al., 2007).

More precisely, the expected weight w_{ir} from school i to school r is

$$W_r \frac{V_i}{\sum_k V_k} \tag{A.1}$$

where $W_r = \sum_i w_{ir}$ is the sum of inflow weights of school r, or total number of students that enrolled in r, and $V_i = \sum_r w_{ir}$ is the sum of outflow weights of school i, or the total number of students that were assigned to i but enrolled in another school. The expected value for $w_{ir}w_{jr}$ becomes

$$W_r^2 \frac{V_i V_{ij}}{\left(\sum_k V_k\right)^2}. \tag{A.2}$$

Summing this over all r and using $\sum_k V_k = \sum_r W_r$, we find the expected value of $\vec{w}_i^{out} \cdot \vec{w}_j^{out} =$

$$\frac{\sum_r (W_r^2)}{\left(\sum_r W\right)^2} V_i V_{ij}. \tag{A.3}$$

Summing the difference between this expected value and the actual values over all pairs ij within an outflow subdistrict s, and then summing over the contributions of every subdistrict s forms the basis for the final positionality expression displayed in equation (A4). We also divide by $\sum_r \left(W_r^2\right)$ to normalize so that outflow positionality is 1 if all senders belong to one module

$$\Pi^{out}\left(P\right) = \sum_{s=1}^{N_s} \left(\frac{\sum_{i \neq j \in s} \sum_r w_{ir} w_{jr}}{\sum_r \left(W_r^2\right)} \right) - \frac{\sum_{i \neq j \in s} V_i V_j}{\left(\sum_r W_r\right)}. \tag{A.4}$$

This definition of positionality is equivalent to the modularity function for weighted bipartite networks in Stouffer et al. (2012). We then use simulated annealing to identify partitions that maximize equation (A4), as well as a similar expression for inflow positionality, $\Pi^{in}(P)$. Inflow positionality has the exact same form with inflow weights instead of outflow weights.

Having identified a set of groups of similar outflow positionality and a set of groups with similar inflow positionality, we then take the intersection of the inflow and outflow groups to find modules of schools that are equivalent in both. For illustration purposes, suppose we have five schools—A, B, C, D, and E. Suppose further that after (separately) maximizing inflow positionality and outflow positionality we find two inflow groups—"IN1" composed of schools A, B, and C, and "IN2" composed of schools D and E—and two outflow groups—"OUT1" composed of A, D, E, and "OUT2" composed of schools B and C. Taking the intersection yields the following three modules of schools that have the same inflow and outflow positionality: "IN1/OUT1" composed of school A, "IN1/OUT2" composed of schools B and C, "IN2/OUT1" composed of schools D and E. Intersecting all the inflow and outflow groups in this manner leads to the emergent subdistricts identified in the paper. Calculating the flows between them, and normalizing those flows with respect to the flows expected by chance alone, yields the map of flow patterns illustrated in Figure 3.