STOCKHOLM UNIVERSITY

Dept of Sociology, Demography Unit / www.suda.su.se

School Choice and Increasing Performance Difference: A Counterfactual Approach

by John Östh, Eva Andersson and Bo Malmberg

Stockholm Research Reports in Demography 2010:11

School Choice and Increasing Performance Difference: A Counterfactual Approach

by

John Östh[*,], Eva Andersson[***] and Bo Malmberg[**]**

Abstract: In recent years the introduction of independent schools have increased school choice alternatives in Sweden. Subsequently, a large share of today's students attend schools other than the school closest to home. Since the compulsory school system is designed to be free of charge and to offer the same standard of education everywhere, increasing school choice opportunities should, hypothetically, not increase the between-school variation in grades. The aim of this paper is therefore to explore to what extent a measured increase in between-school variance in student performance is an effect of school choice. In this paper, the variance in grades of all Sweden resident students graduating from secondary school in the years 2000, 2003 and 2006, is analyzed. Using a counterfactual approach, we compare the students' variation in grades between observed schools of graduation and hypothetical schools of graduation, the latter being constructed by assigning students to their close to home school. The multilevel results indicate that school choice seem to increase between-school variation of grades.

***Department of Social and Economic Geography, Uppsala University | john.osth@kultgeog.uu.se**
**** Department of Human Geography, Stockholm University | bo.malmberg@humangeo.su.se**
***** The Institute for Housing and Urban Research, Uppsala University | eva.andersson@ibf.uu.se**

Introduction

In OECD's Programme for International Student Assessment (PISA), a critical measure of the quality of national school-systems is the between-school variance in student performance. A high between-school variance implies that which school students attend becomes an important determinant of their performance. High levels of between-school variance are problematic for both equity and efficiency reasons (Björklund *et. Al.,* 2004). A high variance indicates large differences in school quality and, hence, students that attending the *wrong* school will be punished by lower grades. Moreover, in PISA, countries with high levels of between-school variance tend to do less well in terms of overall student performance, see Figure 1. As discussed in Werfhorst & Mijs, (2010), this is to be expected if there is an upper limit to individual performance.

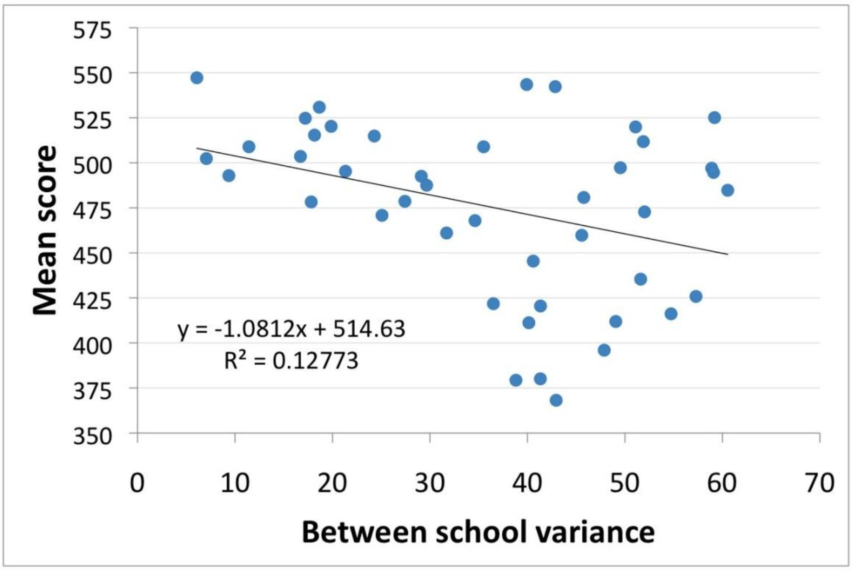

Figure 1 Between school variance and mean student performance in OECD/Pisa data, 43 countries 2000-06. Robust standard error for the slope parameter is 0.358 (Huber–White).

During the past two decades the introduction of independent schools and school choice has changed the Swedish educational system. Until 1992, all comprehensive education, with very few exceptions, took place in public schools. Swedish municipalities had a great deal of self-governance, being able to locate and relocate groups of students to schools on the basis of the municipality's demographic situation, and or to create a mix of rural/urban children in the same school. However, proximity to school was the prevailing ideal in strategies for allocating students to schools. Consequently, each school's catchment area has varied over time, but neither school nor parents have had little influence over the catchment area.

In 1992, independent schools were first introduced in Sweden. While all education remained free of charge for the students, the introduction of independent schools meant that public and non public (independent) schooling agents were paid for each student entering their school. The transition from an almost entirely public schooling system to a system containing several schooling agents spurred competition between schools. In the early years, the few independent schools available played a minor role in Swedish schooling. According to the National Agency for Education, 3.2% of the secondary school graduates attended independent schools in year 2000. This percentage increased to 6.3% in 2003 and 8.5% in 2006 (Skolverket, 2010). While the percentages of independent school attendees remain relatively low, the choices made between public schools are non-negligible. The over-time increase of independent schools, and subsequent loss of students in public schools, has triggered fierce competition between schools (By, 2005). In order to attract students, a majority of the independent schools advertise themselves as pedagogically profiled. The profiling includes; focusing on specific school subjects (sports, language, etc.), using specific pedagogical methods (such as Montessori), or adhering to particular confessional orientations (Friskolornas Riksförbund 2010). As a response, many public schools also make use of pedagogical profiles to attract students. Common profiles among public schools include music-, drama- or sports-orientation. As a result of this increasing competition, parents and students receive advertisement from the different schools (public and independent schools), and parents fill in forms of school-choice. These choices typically are made when the child is about to start first grade, or when the student is about to leave the third, sixth or ninth grade.

The greatest number of independent schools, and of students attending them can be found on upper secondary school level. However, since the upper secondary educations are non-mandatory and the types of educations offered at various schools differ substantially, any analysis of differences in grades between close to home schools and observed schools on this level must be considered fruitless. Primary and secondary school on the other hand, is mandatory and since the very fabric of compulsory school is built on the assumption that all students are being offered equal education opportunities (Education act 2000), a study on the effects of school choice in compulsory school is not only possible but of societal importance.

The importance attached to the goal of equal educational opportunities is reflected in the small differences in quality between Swedish schools in 2000 (OECD, 2001). In reading tests, for example, between-school variance in performance was below 10%, compared to a mean of 36% in the OECD sample. However, as shown in Figure 2, between-school variance in student performance has been rising rapidly since 2000. According to the results of the PISA study, this increase in variance has been paralleled by a decline in mean student performance in science, reading and maths.

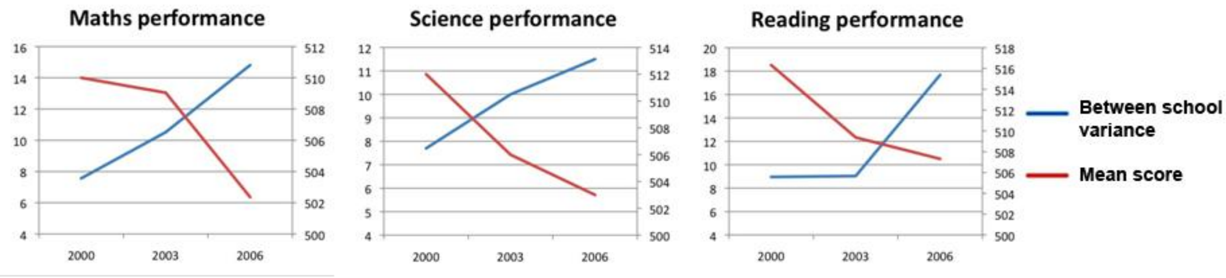

Figure 2 Mean and between School Variance in Student performance, Sweden 2000-2006

The aim of this paper is to explore to what extent this increase in between-school variance in student performance is an effect of school choice. In the Swedish context, two different arguments concerning the effects of school-choice on between school differences can be found. The first argument is that school choice encourages ethnical Swedes to avoid minority-dominated low performing schools (see e.g. Bunar, 2010a; 2010b) and, thus, reinforce school segregation. The opposite argument is that increasing residential segregation is the driving force behind increasing ethnic and socio-economic school segregation (Lindbom & Almgren 2007). Lindbom and Almgren show that particularly ethnic residential segregation increased strongly between 1990 and 2004. Following from this, it is argued that in segregated areas, school choice can actually alleviate the effects of residential segregation. In our study, we will approach this question by estimating between-school variance on performance both for observed schools of graduation, and for hypothetical, close to home schools to which students have been grouped based on home address. If, as suggested, residential segregation, as opposed to school choice, is the driving force behind increasing performance gaps between schools, then increasing between school variance should be identified both in observed schools of graduation, and in the hypothetical close to home schools (the counterfactual situation).

Earlier research

To provide a background to our study we will begin with earlier research, in which a school choice situation is compared with a non-choice situation. Secondly, we will refer to studies relevant to the Swedish context, where school and neighbourhood effects on educational outcome are compared.

Lastly, the motivation of studying school effects during different school regimes, of for example choice and non-choice, is demonstrated through studies of stated school effects.

1. School choice situation compared to a proximity counterfactual situation

Internationally, there are important contributions to the analysis of school choice and market based educational systems (Alegre and Ferrer 2010, Taylor 2001, Reay 2004). There are also geographers in the fields of school choice research, arguing for the importance of race rather than class in the process of choice, (Byrne, 2009; Garcia, 2008) and scholars adding multilevel methods (Harris et al., 2007). Concerning school choice there is an interest in parents choosing to avoid schools with an overrepresentation of students with a foreign background (Saporito, Lareau 1999; Noreisch 2007) and/or choosing between public and private alternatives (Seppanen 2003; Karsten et al., 2003). In addition researchers have constructed counterfactual studies within a choice situation, which is of special interest for our findings (Allen, 2007; Bernelius, 2005; Burgess *et al.*, 2007; Lindbom and Almgren, 2007).

Since the Finnish and Swedish school systems are relatively similar, there is reason to look closer into the results by Bernelius (2005). Bernelius' study on Helsinki provides strong evidence for school choice contributing to increasing differentiation in the schools' student composition. School choice favours schools with larger shares of students from privileged backgrounds, and students choosing a school outside their own catchment area have a higher than average socio-economic family status. The educational outcome is also significantly better in the group choosing a school outside their own catchment area. The method of assessing the impact of school choice in Helsinki was building on individual-level data on each student's elementary school (grades 1 to 6) on their current secondary school (grades 7 to 9) and their educational outcomes in the beginning of the 7th grade.

In a Swedish study by Lindbom and Almgren (2007) they investigate the difference between observed and close to home schools concerning *student composition* in the school area. Lindblom and Almgren find an increase in ethnical segregation and no particular change of socio-economic aspects in school areas; however they do not analyse the students' achievements in association with the segregation increase, that is, the consequences of school segregation. Such an analysis, argue Lindbom and Almgren, requires a more complete data material on schools instead of areas surrounding schools and for cities other than the three largest in Sweden.

In her study of the English secondary school system, Allen (2007) finds that around 50 per cent of the children did not attend their proximity allocation school. Allen estimates that only one in five students is potentially active in sorting between non-faith comprehensive schools, and moreover that school segregation is almost always lower in the close to home counterfactual school than in the observed school. Allen's results further confirms that the social and ability segregation is greater when students sort themselves into schools, and that the difference between school and residential segregation is greatest in the urban areas where students can choose between several different kind of schools. Allens' study is comparable to ours in that it aims to illustrate how changes in the school system have altered the allocation of students to schools.

In addition to these quantitative studies using a counterfactual approach there are studies where the school system, or the school regime as such, is analysed and consequences of sorting and choice discussed. On the basis of PISA results, Alegre and Ferrer (2009) state that market oriented school regimes increase schools social segregation. The study relates to ours on a general level, illustrating how changes in the system affect allocation of students, and in many cases increases differences between schools.

2. School and neighbourhood effects compared

Using a multilevel analysis some researchers have compared the school effects with the neighbourhood effects, an approach which is close to comparing a school choice situation with a non-school choice situation. Neighbourhood is then considered 'the same' as the proximity school concerning composition characteristics. However the approach differs from ours in that they do not use close to home school but instead the residential area. There is, to our knowledge, only two Swedish studies comparing school and neighbourhood effects (Bergsten 2010, Brännström, 2008), but for our purposes a study of Helsinki is also important (Kauppinen 2008).

Brännström analyses average upper secondary final grades in Stockholm 2004, examining school and neighbourhood effects in the same multi level model (cross-classified model). He finds this necessary since estimated effects of concentrations of (dis)advantage and immigrant share at the neighbourhood and school level point in different directions. Being a first generation immigrant in an immigrant dense school is an advantage.

The contexts on both school and neighbourhood levels were characterised by social assistance recipiency, university degree and immigrant density (proportions). Brännström concludes that characteristics attributed to upper secondary schools are more important for educational achievement than neighbourhoods. School effects were around 23 per cent and neighbourhood effects around 3 per cent (compare empty model in this paper). When controlling for individual variables the estimate decreased considerably (school effect around 7 percent) and this Brännström finds, is mainly due to selection bias in the assignment process. The remaining neighbourhood effect when controlling for individual variables is around 2 per cent. (Brännström, 2008)

Bergsten (2010) also compares school level effects on education with neighbourhood level effects using a multilevel method. Her approach does not exactly correspond to ours, but is similar enough to take into consideration. (We do compare observed school population with a counterfactual close to home school population and, as stated earlier, we consider close to home school almost the same as neighbourhood concerning characteristics.) In accordance with other studies comparing neighbourhood and school level Bergsten finds greater effects from schools than neighbourhoods (empty model, software MLwiN) on students upper secondary education. In her models the school context explain about 21% of the total variance in grades among upper secondary school students, whereas the neighbourhood level explain about 4% (Bergsten 2010, p. 229). Part of the explanation of this greater school effect, according to Bergsten, is the possibility of indirect effects from the neighbourhood (Leckie, 2009). Upper secondary education might well be influenced by earlier results from primary school which is very much correlated to the population in the neighbourhood.

Similar to Brännström (2008) and Bergsten (2010) researchers on educational opportunities and neighbourhood effects, increasingly acknowledge the importance of schools as mediating the effects from neighbourhoods. Some are admitting greater importance to the school in shaping the future achievements of students. In his study on secondary education in Helsinki Kauppinen (2008) finds that the effect on neighbourhood's educational composition on educational choice is predominantly mediated by the school's socio-economic composition (Kauppinen, 2008).

There are a handful of multilevel studies of school effects in Sweden and also their purpose has been different, see Sellstrom and Bremberg (2006) for problems with comparing multilevel studies. Dryler's study from 2001 concerns ethnic segregation between schools and whether or not it has effect on youths' grades, uncompleted grades and the probability of passing on to theoretical programs in upper secondary education. Dryler finds positive effects of percentage foreign born on educational outcomes in Sweden; we will need to consider this controversial result carefully. She finds that large proportions of foreign born affected average grades negatively in 1990/91. Interestingly these differences had disappeared for students' average grades in 1997 (compare positive effects of ethnic composition on upper secondary education grades in Norway, (Fekjaer and Birkelund, 2007)).

Szulkin and Jonsson (2007) find depressed grades among first generation immigrants in Sweden, and also that first generation immigrants in a school tend to depress grades among students in general. A last example on Swedish school effects (multilevel) also show the result that individuals and their households did not fully explain the school variation; a report from Statistics Sweden (SCB, 2007). That is, multilevel modelling is motivated and there is reason to explore differences in school allocation during different school regimes (i.e., choice, non-choice). In the SCB study several variables are controlled for, including parental education, family composition, ethnic background, time in Sweden and the type of housing. School effects remain after controlling for both neighbourhood and school composition. The highest risk of ineligibility to upper secondary school is found among adolescents in immigrant dense areas who attended a school with high shares of students with foreign background (SCB, 2007, compare Andersson, Östh and Malmberg 2010).

The National Agency for Education analyses school level effects with respect to both socioeconomic and migration background and detect a substantial increase over the time period 1998-2004 (Skolverket, 2006). Consequently, the opportunity for students to learn will not be equal and will, to a larger extent, depend on which school the students attend according to the National Agency for Education. The study is based on students' grades and also using PISA data. To conclude there are known school effects in Sweden, but what will a comparison between a choice and a non-choice situation show in the matter of school effects? In the following we will explore this question.

Research design

In our study, we will approach the question of increased between school variance by estimating students' performance both for observed schools of graduation, and for counterfactual, close to home schools to which we grouped students based on home address. If residential segregation is the driving force behind increasing performance gaps between schools, then increasing between school variance should be observed both in schools of graduation, and in the hypothetical counterfactual close to home schools. If however school choice is a more important driving force behind the increasing between-school variance in students' performance, increasing between school variations should be observed primarily in the observed schools of graduation. While the direction and strength of the coefficients in our models merits attention, the key focus in this study is rather to assess the overall effects on students' performance as a consequence of a school choice situation.

In our study, all graduating secondary school students in Sweden in the years 2000, 2003 and 2006 are being included. The inclusion of all students, rather than a sample of students, is necessary in order to create a counterfactual analysis. However, the use of a total population does create a potential geographical problem of choice that needs to be addressed. Since almost all schools are located in urban areas while students reside in both urban and rural areas, a geographical mismatch in choice of schools seems unavoidable. Students in smaller urban areas and in rural districts have longer distances and fewer schools to choose between. The problem is however reduced due to two factors. First, a school bussing system shuttles students from smaller urban and rural areas in to the larger urban areas free of charge, reducing the geographical problem of choice. Second, since our key aim is to determine whether there is a significant difference in the variation of grades between the observed and counterfactual schools of graduation, a reduction in the number of available schools to choose from in rural and smaller urban areas, do not risk to overestimate the effects of school choice on the variation in grades.

To assess the effects of school choice on students' performance, each student needs to be assigned to two schools, using two different methods. Firstly, the observed school of graduation, straightforwardly found in our data, is used. The second counterfactual assignment procedure is more complex. In the counterfactual procedure the school closest to home needs to be assigned to each student. Schools differ significantly in size and consequently, a small school cannot assign all neighbouring students, simply because it is located closer to home than all other school alternatives. To solve the school size constraint, a computer program has been developed, wherein the coordinates of all students and all schools are included. Each school is also given a maximum capacity value, set to be equal to the number of observed students graduating from the school. Starting at the coordinates of each school, students are assigned to the school on basis of proximity. In the first round, all students living within 50 meters from the schools are assigned to the schools, in the second round all students within 100 meters, thereafter 150 meters, 200 meters and so on. Whenever a school reaches their student max capacity, no further student is assigned to that school. Remaining students are assigned to other relatively close to home schools, still having vacancies. The school-matching iterations continue until all students have been assigned to a school. Since the observed schools of graduation, with very few exceptions, are located in the municipalities where the students reside, the total schooling capacity in any municipality is quintessentially the same as the total amount of students. Consequently, assigning students to schools is a fairly swift procedure in which more than 99% of the students are assigned to either their closest, second closest or third closest to home- school. All students are assigned to schools to which they would have been able to travel on a daily basis. The principle behind the creation of close to home schools is illustrated in Figure 3 below. The upper parts of the figure illustrate three different phases in the construction of school catchment areas. In the lower part a close-up illustration of the 50-meter circles is illustrated. By using the two different ways of assigning students to schools, each student is given a code representing the observed school of graduation, and a code representing the counterfactual school of graduation (close to home). In cases where the student graduates from the close to home school, the observed and counterfactual school will be the same; in all other instances the observed and counterfactual (close to home) schools will be different.

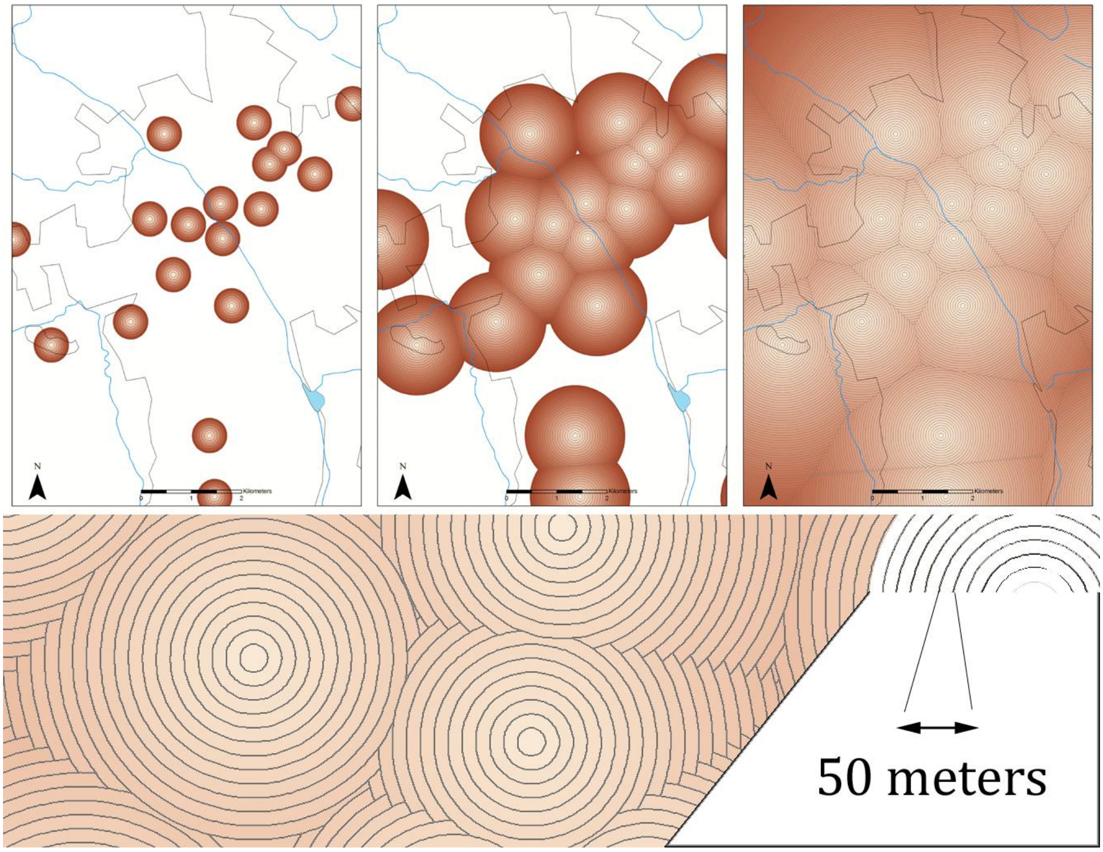

Figure 3 Illustration of the principle behind assignment-areas in the creation of close-to-home schools. The central-most part of each circle represents the location of a school. Additional circles represent enlargements in the search-areas by 50 meters. As long as the school has not reach its capacity maximum, all students within the catchment area is assigned to the school.

In order to assess significant differences in the variance of grades due to school choice, multilevel models were designed (using MLwiN). Multilevel models are commonly used in studies wherein the explanatory power is assumed to differ between hierarchical levels (Raudenbush and Willms 1995, Subramanian 2004). The performance of students is such an example. The performance of a student can partly be attributable to the student's own capacity, his/her environment, parents etc. However, the performance of the student may also partly be explained by the school context in which he or she studies. In order to better estimate the performance of students, the analysis of the student's performance must encompass the underlying nested hierarchical structure. In this study, two hierarchies are used: the individual level and the school level (see Table 2 for number of students and schools each year). The results from the multilevel analysis will separate to what extent the individual and the school respectively affects the performance of students.

The analysis is be conducted in three steps; an empty model, a student-level model and finally a student- and school-level model. In the first step, an empty model is designed for each year, and for each of the two datasets (observed school and counterfactual school). The empty model contains no explanatory variables and consequently the variation in grades between students and schools is due to difference in contribution to the explanation at the different levels. As such a possible variation in grades in either of the two levels can be used as benchmarks in comparisons to more complex models, having several explanatory variables.

In the second step, student-level models are designed. In addition to the empty model, the student-level models contain five student-level variables. Finally, in the third step student- and school-level models are designed. In addition to the five variables introduced in the student-level model, these models also include five school-level variables. The formulation of a multi(in this case two)-level analysis of variance is as follows in equations 1-4.

Eq.1 $$y_{ij} = \beta_{0ij}x_0 + \beta_{1ij}x_1 + ...\beta_{nij}x_n$$

Wherein y_{ij} represents grade, $\beta_{0ij}x_0$ represents intercept and error terms, and $\beta_{nij}x_n$ represents fixed parameters in the regressions (the independent variables are presented below). With the exception of the empty model, which only includes β_{0ij}, consisting of intercept, and error terms/variation on school and student levels (Eq. 2).

Eq. 2 $$y_{ij} = \beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$$

The unobserved variation in grades is categorized in to school level variation *u,*

Eq. 3 $$u_{0j} \sim N(0, \sigma_{u0}^2)$$

and individual level variation *e.*

Eq. 4 $$e_{0ij} \sim N(0, \sigma_{e0}^2)$$

We run all regressions using two different populations. First, all students graduating from schools with a minimum of 15 graduating students; secondly, all students regardless of the size of the school of graduation. In order to reduce the potential inflation in variance due to small schools, all analyses presented in the results section make use of the population containing at least 15 graduating students per school ("Min 15" in table below). In the appendix section, additional tables illustrating the results of analyses wherein all schools/students are included ("All" in table below).

Table 1 Structure of the statistical analyses in the article. Analyses * are presented in appendix.

	Observed schools						Counterfactual schools					
Models	**Year 2000**		**Year 2003**		**Year 2006**		**Year 2000**		**Year 2003**		**Year 2006**	
Empty model	Min 15	All*	Min 15	All*	Min 15	All*	Min 15	All*	Min 15	All*	Min 15	All*
Student-level model	Min 15	All*	Min 15	All*	Min 15	All*	Min 15	All*	Min 15	All*	Min 15	All*
Student- and school-level model	Min 15	All*	Min 15	All*	Min 15	All*	Min 15	All*	Min 15	All*	Min 15	All*

All data material used in this study is collected from the PLACE-database (Department of Social and Economic Geography, Uppsala University), see Table 2. The database contains detailed annual information about all Sweden resident individuals' work-related, socioeconomic and educational status during the years 1990 through 2006. The material in this study makes use of the following variables presented in Table 2.

The residential coordinates of all students graduating from secondary schools in the years 2000, 2003 and 2006 were selected. The coordinates are arranged on a hundred meter level, allowing for very precise geographical analyses. The school coordinates represent the location of all schools for graduating secondary students in the years 2000, 2003 and 2006. Students' graduating grades from secondary school are used to test the variance in grade variations between students and schools in 2000, 2003, and 2006. The same grading system has been active throughout the years in the study. The grades describe the sum of grades from different subjects. The following subject grades are given; zero for fail, ten for pass, fifteen for pass with distinction, and twenty for pass with special distinction. The average student achieves a final grade (sum of subject grades) of approximately 200 points, however, slight differences in the mean values over time are detectable (less than one point). In order to increase the validity of over-time comparisons, the grades are Z-score transformed, assuring that the average grade is similar throughout the studied years. As illustrated in Figure 4 below, the grades are not perfectly, normally distributed; however, the pattern of distribution of grades is reasonably similar in all of the included years. A noteworthy deviation is the increasing over-time standard deviation of grades; in the figure below, the deviation can be detected visibly, especially among the failing and the top students. Expressed in numbers, the standard deviation of the observed grades (untransformed, mean of 200) rose from 60.5 in year 2000, to 66.4 in 2003 and 66.8 in 2006.

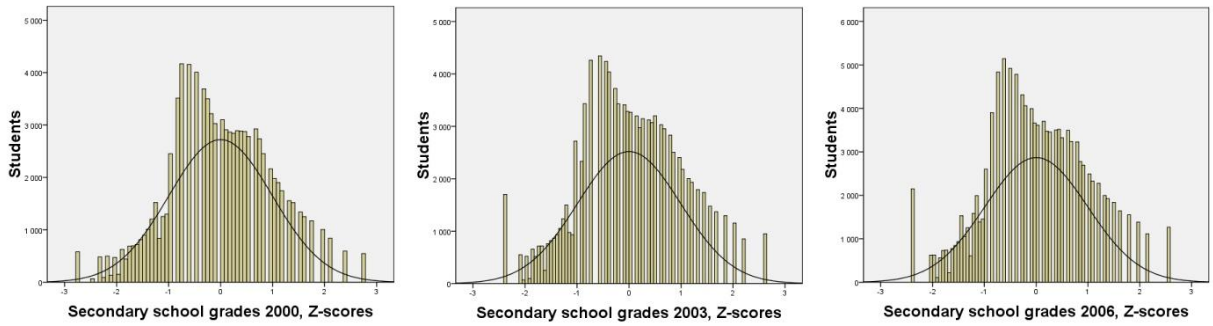

Figure 4 Illustrate three histograms of students' grades for the years 2000, 2003, 2006. Number of students is illustrated on the y-axis and the Z-score of final grades on the x-axis.

Measuring variation in grades for observed schools of graduation is relatively straight forward. However, measuring variation for counterfactual close to home schools of graduation is more complex. Students' observed grades are to a large extent a result of the individual student's ability, ambition and socioeconomic background, but also a result of teachers' skills, and abilities and opinions towards education among fellow students. Using the observed grades in the counterfactual close to home schools means that we implicitly include effects that are endogenous to the observed school of graduation. In order to cater for this problem, we make use of the grade variable wherein the observed school effect is subtracted from each student's grade. In the results section, the effects of using both types of grades for the regression results on counterfactual schools are being discussed further.

Beside variables depicting grades and schools, five variables, representing the students' socioeconomic background are included as well, see Table 2. The students' socioeconomic status variables include firstly the variable born abroad. The variable distinguishes between Sweden born individuals and individuals born in any other country. Secondly, the sex variable distinguishes between boys and girls. Thirdly, parents' education show if any of the student's parents have a post upper-secondary education. The fourth variable confirms if the student's household received social benefits during the year of graduation from secondary school. Lastly, the variable student's disposable income represents the log of hundreds of Swedish kronor (SEK) disposable to the student during the year of graduation. The income is calculated using a child-rearing cost index times the sum of the annual household income. The variable is constructed and delivered by Statistics Sweden (SCB). The school-variables are constructed using the corresponding student variables. With the exception of the student's disposable income variable, all school-level variables are constructed to represent the share of students in each school being either foreign born, female, having higher educated parents or residing in households dependant on social benefits. The student's disposable income variable on school level is the mean disposable income among the students in each school. The school level variables are calculated separately for each of the two datasets (observed and counterfactual).

Table 2 Variables in analysis, data from PLACE. (Number of students in models with a minimum of 15 students per school).

Level of analysis	Variables	Description
1. Individual	Student no.	Count (94 530 in year 2000, 105 899 in 2003, 120 472 in 2006)
	Born abroad	Sweden born (0) born in any other country (1)
	Sex	1= female 0=male
	Parents' education	If any of student's parents have post upper-secondary education, the parents' education variable value equals 1. In all other cases, 0.
	Social benefit	If student's household receives social benefit during the year of graduation from secondary school, the variable social benefit equals 1, in all other cases, 0.
	Student's disposable income	Variable representing the log of hundreds of SEK disposable to the student during the year of graduation. Income value is calculated using a child-rearing cost index times the sum of the household income.
2. School	School no.	Count (1 088 in year 2000, 1 254 in 2003, 1 194 in 2006)
	Born abroad	Share of students born abroad in each school
	Sex	Share female students in each school
	Parents' education	Share students in school with parents having a post upper-secondary education
	Social benefit	Share of students' households with social benefit in each of the schools
	Student's disposable income	Mean student disposable income (log of hundreds of SEK) in each of the schools.
Dependent variable	Grades	Sum of grades from different subjects. (Z-score of grades, annual overall mean = 0)

Results

The results section contains two parts. In part one, maps are used to illustrate the relationship between school choice and students' residential locations in a Swedish city. In part two, multi-level regressions are run containing all graduating students and schools of graduation in Sweden for the years 2000, 2003 and 2006.

Maps

An example of the geographical pattern of school choice in a Swedish city is shown in Figure 5 below. The figure contains the locations of schools and students during the last year of compulsory school (year 9) in the mid-sized city of Uppsala (approximately 120 000 inhabitants in the urban area). Since the level of geography in the figures is relatively detailed, no reference to the year will be given due to confidentiality. In Figure 5, the student catchment areas of two independent schools (Map 1-2) and five public schools are illustrated (Map 3-7). Smaller dots represent the residential locations of the ninth grade secondary school students, while the larger dots represent the locations of schools. As illustrated in Map 3-7, the majority of the public schools' students are enrolled from the school's neighbourhood. However, the student catchment areas of the independent schools are much wider, attracting students from large parts of the city. In addition, there are student clustering patterns in Map 1. The independent school attracts many students from the southern-most parts of Uppsala, located far from the school itself. The clustering of students in the southern part of the city may point towards a disturbing segregation pattern. The southern middle-class residential area has its close to home school in the neighbouring lower-class and immigrant-dense area to the north-west. Choosing an alternative school in the city centre may be a strategy to avoid a low-status school.

Map 8 in Figure 5 illustrates the share of local students studying at each school. The size of the hexagons illustrates the share of locally assigned students. The larger the size of the hexagon, the larger is the share of locally assigned students and vice versa. The pattern reveals that city-centre schools attract students from many different locations, while schools in the fringe of the urban area attract local students.

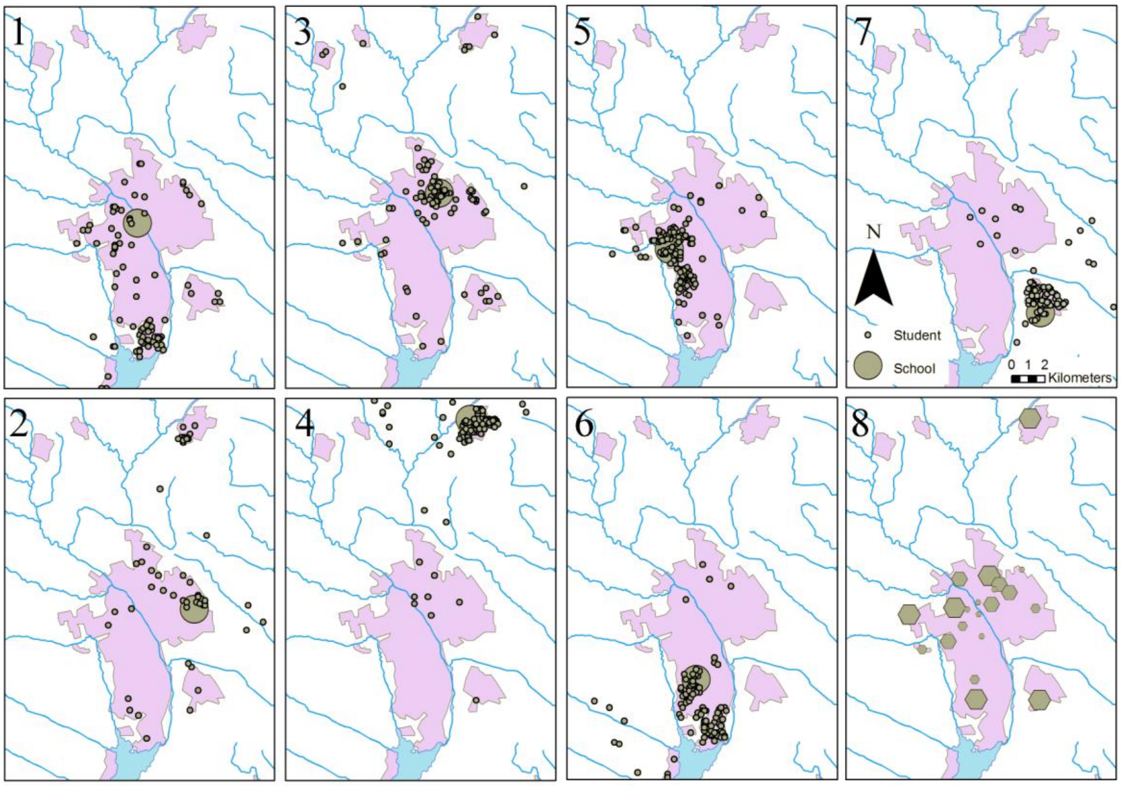

Figure 5 Maps 1 and 2 illustrate the distribution of two different independent schools and their students. Map 3 to 7 illustrate the distribution of students attending public schools in Uppsala. Map 8 illustrates the distribution of all secondary schools in the Uppsala area.

Multilevel analyses

Multilevel analyses are conducted on two different datasets, observed and counterfactual, during three different years: 2000, 2003 and 2006. Three different models are used in these analyses; 1) an empty model, 2) a student-level model and 3) a student- and school-level model.

In Table 3, the results from the empty-model regressions are illustrated. The results point to notable differences between the observed and counterfactual datasets. The results indicate that the variance in grades among the counterfactuals is largely attributable to the student; consequently the school level variance is small, ranging from 4.80% to 5.53%. Importantly, however, there is a variance at the school level motivating the use of multilevel models. The pattern for the observed dataset is different. Besides displaying significantly greater school variance than noted for the counterfactuals, the school variance is also becoming more pronounced over time. The strong increase in variation in 2003 coincide with the increase of independent schools in Sweden, and more importantly, the public schools response to increased competition through specialization and profiling. Hence the empty model indicates that school choice (demonstrated in the observed dataset) and not residential segregation (demonstrated in the counterfactual dataset) is the factor that determines the greatest share of variance.

Table 3 Empty model

	Observed school			Counterfactual school		
	2000	**2003**	**2006**	**2000**	**2003**	**2006**
FIXED	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Intercept	-0,009 (0,009)	-0,018 (0,011)	-0,017(0,011)	-0,005(0,008)	0,015(0,008)	0,017(0,008)
RANDOM	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation
Variance components						
Student level	93,94%	87,35%	86,51%	95,20%	94,47%	94,57%
	0,930 (0,004)	0,891 (0,004)	0,885 (0,004)	1,011(0,005)	1,094(0,005)	1,114(0,005)
School level	6,06%	12,65%	13,49%	4,80%	5,53%	5,43%
	0,060 (0,003)	0,129 (0,006)	0,138 (0,006)	0,051(0,003)	0,064(0,003)	0,064(0,003)
Sum	100%	100%	100%	100%	100%	100%
-2*log likelihood	263304,6	291239,8	330273,4	270591,6	311818	355977,5

In Table 4 estimates from the student-level model including five fixed student level variables are illustrated. The results suggest that by including student level variables, approximately 15%-20% of the variations in both the observed and counterfactual datasets are explained (81-85% unexplained as noted in Table 4). However, the observed patterns from the empty models remain. The school component in the counterfactual dataset explains only approximately 3 percent of the variation in all of the years while the school components in the observed dataset ranges from approximately 3 % in 2000 to 11% in 2006. The results suggest that even though student-level effects are controlled for, schools rather than neighbourhoods (counterfactual dataset) determine a greater share of the variation in grades (compare Burgess *et al*., 2007).

The fixed student-level variables included in the models in Table 4, all have significant and anticipated directions in line with earlier studies (Andersson, Subramanian 2006). Being born abroad and residing in a home dependent on social benefits are the two variables that are negatively associated with grades. Of the two variables, social benefits dependency has the greatest negative impact on students' grades. In contrast, being a girl and/or a child to parents with higher education is strongly and significantly associated with higher grades. The disposable income level is positively associated with higher grades, however less strong. The similarities in direction and significance of the fixed student variables, between the two datasets, and over time, indicate that the included student-level variables are relatively stable.

Table 4 Student level model.

	Observed school			Counterfactual school		
	2000	2003	2006	2000	2003	2006
FIXED	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Intercept	-0,837 (0,020)	-0,747 (0,019)	-0,859 (0,018)	-0,848(0,020)	-0,72(0,019)	-0,822(0,018)
Student variables						
Born abroad	-0,080 (0,009)	-0,082 (0,009)	-0,094 (0,008)	-0,078(0,009)	-0,067(0,009)	-0,069(0,009)
Female	0,408 (0,006)	0,359 (0,005)	0,380 (0,005)	0,410(0,006)	0,363(0,006)	0,382(0,006)
Parents' education	0,601 (0,006)	0,556 (0,006)	0,551 (0,005)	0,618(0,006)	0,591(0,006)	0,586(0,006)
Social benefit	-0,604 (0,011)	-0,688 (0,011)	-0,656 (0,011)	-0,610(0,012)	-0,711(0,013)	-0,676(0,013)
Student's disp. Income	0,052 (0,002)	0,045 (0,002)	0,056 (0,002)	0,052(0,002)	0,043(0,002)	0,053(0,002)
RANDOM	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation
Variance components						
Student level	77,88%	73,63%	72,73%	79,10%	81,26%	81,58%
	0,771 (0,004)	0,751 (0,003)	0,744 (0,003)	0,840(0,004)	0,941(0,004)	0,961(0,004)
School level	3,23%	10,59%	11,34%	2,92%	3,37%	3,65%
	0,032 (0,002)	0,108 (0,005)	0,116 (0,005)	0,031(0,002)	0,039(0,002)	0,043(0,002)
Sum	81,11%	84,22%	84,07%	82,02%	84,63%	85,23%
-2*log likelihood	244 913,70	272 805,70	308523,30	252918,3	295559,4	337990,2

In Table 5 estimates from the student- and school-level model including five fixed student level and five fixed school-level variables are illustrated. The results in Table 5 clearly indicate that the introduction of fixed school level variables affects the variance components to a small extent. This is evident when comparing the relative small reduction of -2*log likelihood in Table 5 with the corresponding -2*log likelihood values in Table 4. The small but noticeable reduction in school level variance (app. 0.3%-0.6%) is similar in both of the datasets. In the counterfactual dataset, the variance attributable to the school decreases to less than three percent, while the percentage values in the observed dataset ranges between 2.63 in 2000 to 11.05 in 2006. The results suggest that while almost all school-related variance is explained in the counterfactual dataset, much variance is unaccounted for in the observed dataset; especially in years 2003 and 2006.

The additional variables in Table 5 comprise five school-level variables. The school-level variables are essentially aggregate student-level variables, depicting the proportion of the various groups of students on each school. School-level variables point at the social and economic context of the student. The results indicate that parents' education is positively and significantly correlated with higher grades on both student- and school level in all of the regressions. Similarly, social benefit is negatively correlated with grades on both student- and school level. Remaining three variables are less straightforwardly interpreted; being a girl is strongly and significantly associated with higher grades on the student level. On school-level however, shares of girls render insignificant results. The disposable income variable renders mixed but relatively weak results. In the counterfactual dataset, the variable provides either insignificant or weak and negative values. In the observed dataset, the results interestingly changes over the years , from almost no effect in 2000, to significantly negative in 2003 and significantly positive in 2006. Finally, the born abroad variable renders mixed results. On the student level, being born abroad is strongly and negatively associated with grades. However, on the aggregate school-level, the pattern is different. In year 2000, both observed and counterfactual datasets display negative correlations between grades and increasing shares of foreign born. However, during 2003 and 2006, a transition from negative to positive values is evident.

The included variables only partly explain the total variance in grades. In the student- and school-level models (Table 5) only between 15 and 20 per cent of the total variation is explained. The unexplained variance on student level can, among other unknown factors, be attributable to the student's intellectual ability, and support from relatives and friends. Due to data limitations, these kind of variables can't be included in our analyses. In the counterfactual datasets, almost all variation is observed on student level; in the observed datasets student level variance is lower, while school level variation is significantly greater than in the counterfactual datasets. The increase in school level variance, and reduction in student level variance, in the observed datasets points to the effects of school choice. The choice driven sorting of students with different abilities, support from relatives and friends together with other unknown factors consequently aggravate differences between schools. Choosing to attend a confessional school, a school with a specific pedagogic profile, or a school with extra music, drama or sports, or simply the school down town, consequently also means choosing a schooling context facilitating or impeding ones educational performance.

As mentioned in the research design section, the observed and counterfactual datasets are regressed using different dependent variables (grades). In the counterfactual datasets, grade variables in which the endogenous effect of the observed school has been eliminated, are being used to reduce bias. Disregarding the problem of endogeneity, using observed grades in regressions on counterfactual datasets, we end up with results in which the fixed parameters on individual and school levels are similar to the results presented in Tables 3-5. The variations in grades, attributable to either schools or students, do however produce interesting results. Using observed grades in the counterfactual regressions reduces variation on school level even further. Throughout the studied period and in all of the models (empty, student level, and student and school level) the variation attributable to schools, drops between 1.5% and 2%. The drop in school level variation potentially points to the scale of the neighbourhood effect. (These results can be obtained from the authors upon request.)

Table 5 student- and school level model

	Observed school			Counterfactual school		
	2000	**2003**	**2006**	**2000**	**2003**	**2006**
FIXED	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Intercept	-1,042 (0,186)	0,063 (0,297)	-1,775 (0,296)	-0,507(0,250)	-1,005(0,230)	-0,475(0,233)
Student variables						
Born abroad	-0,075 (0,009)	-0,082 (0,009)	-0,092 (0,008)	-0,071(0,009)	-0,057(0,009)	-0,066(0,009)
Female	0,408 (0,006)	0,359 (0,005)	0,380 (0,005)	0,410(0,006)	0,364(0,006)	0,382(0,006)
Parents' education	0,593 (0,006)	0,555 (0,006)	0,550 (0,005)	0,612(0,006)	0,583(0,006)	0,579(0,006)
Social benefit	-0,599 (0,011)	-0,685 (0,011)	-0,653 (0,011)	-0,606(0,012)	-0,702(0,013)	-0,671(0,013)
Student's disp. Income	0,052 (0,002)	0,045 (0,002)	0,055 (0,002)	0,053(0,002)	0,043(0,002)	0,053(0,002)
School variables, share of students						
Born abroad	-0,065 (0,059)	0,075 (0,094)	0,234 (0,095)	-0,147(0,071)	-0,008(0,067)	0,098(0,069)
Female	-0,046 (0,098)	0,008 (0,141)	0,162 (0,136)	0,042(0,095)	-0,072(0,090)	-0,087(0,092)
Parents' education	0,473 (0,043)	0,080 (0,072)	0,157 (0,069)	0,565(0,057)	0,731(0,057)	0,762(0,058)
Social benefit	-0,113 (0,113)	-0,665 (0,207)	-0,797 (0,212)	-0,068(0,130)	-0,534(0,135)	-0,462(0,146)
Student's disp. Income	0,006 (0,021)	-0,100 (0,033)	0,095 (0,033)	-0,069(0,029)	0,004(0,026)	-0,078(0,027)
RANDOM	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation
Variance components						
Student level	77,98%	73,63%	72,73%	79,10%	81,26%	81,58%
	0,772 (0,004)	0,751 (0,003)	0,744 (0,003)	0,840(0,004)	0,941(0,004)	0,961(0,004)
School level	2,63%	10,29%	11,05%	2,45%	2,50%	2,80%
	0,026 (0,002)	0,105 (0,005)	0,113 (0,005)	0,026(0,002)	0,029(0,002)	0,033(0,002)
Sum	80,61%	83,92%	83,77%	81,54%	83,77%	84,38%
-2*log likelihood	244 778,30	272 755,50	308490,90	252786,1	295323,8	337789,1

Concluding discussion

The aim of this paper is to explore to what extent the increase in between-school variance in student performance is an effect of school choice. Increasing levels of variation in grades in Sweden are factual and must be viewed as problematic for two reasons in particular. First, since the very fabric of compulsory school is built on the assumption that all students are being offered equal education opportunities. Increasing variance in grades, and increasing differences in variation in grades between schools, threatens the very idea of the Swedish compulsory school system. Second, according to PISA results, countries with high levels of between-school variance tend to do less well in terms of overall student performance.

In the Swedish debate, there have been two different arguments concerning the effects of school-choice on between-school differences in grades. In the first argument, school choice propels increasing between-school variations in grades, since choosing in practice means avoiding schools with poor reputations. In the second argument increasing *residential* segregation drive the increasing between school variations in socioeconomic background and grades. Hence, school choice may be perceived as a viable instrument for the reduction of between-school variations in grades. To test the validity of the two arguments, all graduating ninth grade students in the years 2000, 2003 and 2006 have been included in multilevel analyses. Each student has been assigned to two different datasets 1) the observed school of graduation dataset and 2) the counterfactual (close to home) dataset. If the between-school variation in grades is greater in the observed dataset, the results suggests that choosing schools rather than residing in a segregated area is more important when determining the degree of variation in grades. If however, the between-school variation is greater in the counterfactual dataset, residential segregation must be considered as a more important factor.

The results from the analyses made in this paper confirm that school choice, rather than residential segregation, is a more important factor determining variation in grades than is residential segregation.

The empirical analysis in this paper confirms the PISA-based finding that between-school variance in student performance in the Swedish school system has increased rapidly since 2000. We have also been able to show that this trend towards increasing performance gaps cannot be explained by shifting patterns of residential segregation. A more likely explanation is that increasing possibilities for school choice have triggered a process towards a more unequal school system. A rapid growth in the number of students attending voucher-financed, independent schools has been an important element of this process. Although the voucher system was introduced already in the early 1990s, the substantial increase in independent schools, and the public schools' response to the increasing competition, did not reshape the school system until the early 2000s. That students increasingly have selected non-neighbourhood schools is evidenced by the fact that the median distance from home to schools in our data has increased over the period. In 2000, this distance was 1.6 kilometres. In 2006 it had increased to 1.8 kilometres.

The idea of voucher-based independent school choice is commonly ascribed to Milton Friedman (1955). Friedman's argument was that vouchers would decrease the role of government and expand the opportunities for free enterprise. He also believed that the introduction of competition would lead to improved school results. As we have seen in the Swedish case, this has not happened. As school choice has increased, differences between schools have increased but overall results have gone down. As has proved to be the case with other neo-liberal ideas, school choice—when tested—has not been able to deliver the results promised by theoretical speculation.

Acknowledgements

We gratefully acknowledge financial support from Vetenskapsrådet through the SIMSAM project Register-based Research in Nordic Demography and Vetenskapsrådets utbildningsvetenskapliga kommitté. We thank colleagues at the Stockholm University Demography Unit and The Institute for Housing and Urban Research at Uppsala University for valuable comments and assistance.

References

ALEGRE, M. A. and FERRER, G. (2010) School regimes and education equity: some insights based on PISA 2006, *British Educational Research Journal*, 36, pp. 433 - 461.

ALLEN, R. (2007) Allocating Pupils to Their Nearest Secondary School: The Consequences for Social and Ability Stratification, *Urban Stud*, 44, pp. 751-770.

ANDERSSON, E. and SUBRAMANIAN, S. V. (2006) Explorations of neighborhood and educational outcomes for young Swedes, *Urban Studies*, 43, pp. 2013-2025.

ANDERSSON, E., ÖSTH, J., and MALMBERG, B. (2010) Ethnic segregation and performance inequality in the Swedish school system: A regional perspective, *Environment and Planning A*, Accepted for publication

BERGSTEN, Z. (2010). *Bättre framtidsutsikter? Blandade bostadsområden och grannskapseffekter : En analys av visioner och effekter av blandat boende (Better prospects through social mix? Mixed neighbourhoods and neighbourhood effects : An analysis of the purpose and effects of social mix policy)*, Uppsala: Uppsala University.

BERNELIUS, V. (2005) Onko oppimistulokset valettu betoniin? Tutkimus Helsingin kaupunkirakenteen ja peruskoulujen oppimistulosten yhteydestä ja kouluvalintojen vaikutuksista. (Are educational outcomes real estate? A study on the relationship between urban socio-economic structure and educational outcomes, and about the effect of school choice). , (Translation of content into English provided by the author).

BJÖRKLUND, A., EDIN, P.-A., FREDRIKSSON, P., and KRUEGER, A. (2004). Education, equality and efficiency - An analysis of Swedish school reforms during the 1990s: IFAU - Institute for Labour Market Policy Evaluation.

BRÄNNSTROM, L. (2008) Making Their Mark: The Effects of Neighbourhood and Upper Secondary School on Educational Achievement, Eur Sociol Rev 24(4):463-478.

BUNAR, N. (2010a) The Geographies of Education and Relationships in a Multicultural City, *Acta Sociologica*, 53, pp. 141-159.

BUNAR, N. (2010b) Choosing for quality or inequality: current perspectives on the implementation of school choice policy in Sweden, *Journal of Education Policy*, 25, pp. 1-18.

BURGESS, S., MCCONNELL, B., PROPPER, C., and WILSON, D. (2007). The Impact of School Choice on Sorting by Ability and Socio-economic Factors in English Secondary Education,in *Schools and the Equal Opportunity Problem (CESifo Seminar Series),* ED. Woessmann, L & Peterson P.E., The MIT Press, Cambridge (MA)

BY, U. (2005) En av fem går i skola långt bort (One in five in a school far away). Interview with Nihad Bunar. Newspaper, Dagens Nyheter, 2005-11-14, http://www.dn.se/sthlm/en-av-fem-gar-i-skolalangt-bort-1.410989

BYRNE, B. (2009) Not just class: towards an understanding of the whiteness of middle-class schooling choice, Ethnic and Racial Studies, 32, pp. 424-441.

Education Act (2000) (Translation of Skollagen, 1985:1100, issued 12 December 1985. The translation refers to the Swedish copy from June 2000.) http://www.sweden.gov.se/sb/d/574/a/21538. Stockhom, Ministry of Education and Research

FEKJAER, S. N. and BIRKELUND, G. E. (2007) Does the Ethnic Composition of Upper Secondary Schools Influence Educational Achievement and Attainment? A Multilevel Analysis of the Norwegian Case, *Eur Sociol Rev*, 23, pp. 309-323.

FRISKOLORNAS RIKSFÖRBUND (2010) http://www.friskola.se

GARCIA, D. R. (2008) Academic and Racial Segregation in Charter Schools: Do Parents Sort Students Into Specialized Charter Schools? Education and Urban Society, 40, pp. 590-612.

HARRIS, R., JOHNSTON, R., and BURGESS, S. (2007) Neighborhoods, Ethnicity and School Choice: Developing a Statistical Framework for Geodemographic Analysis, Population Research and Policy Review, 26, pp. 553-579.

KARSTEN, S., LEDOUX, G., ROELEVELD, J., FELIX, C., and ELSHOF, D. (2003) School Choice and Ethnic Segregation Educational Policy, 17, pp. 452-477.

KAUPPINEN, T. (2008) Schools as Mediators of Neighbourhood Effects on Choice Between Vocational and Academic Tracks of Secondary Education in Helsinki, *European Sociological Review* 24, pp. 379-391.

LECKIE, G. (2009). *The complexity of school and neighbourhood effects and movements of pupils on school differences in models of educational achievement* Bristol: Centre for Multilevel Modelling, University of Bristol, UK.

LINDBOM, A. and ALMGREN, E. (2007) Valfrihetens effekter på skolornas elevsammansättning: Skolsegregationen i Sverige, in A. LINDBOM (Ed.) *Friskolorna och framtiden -segregation, kostnader och effektivitet*, pp. 89-119. Stockholm: Institutet för framtidsstudier.

NOREISCH, K. (2007) School catchment area evasion: the case of Berlin, Journal of Education Policy, 22, pp. 69-90.

OECD. (2001). ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT Knowledge and skills for life: first results from the OECD Programme for International Student Assessment executive summary Paris.

RAUDENBUSH, S. W. and WILLMS, J. D. (1995) The Estimation of School Effects, Journal of Educational and Behavioral Statistics, 20, pp. 307-335.

REAY, D. (2004) Exclusivity, Exclusion, and Social Class in Urban Education Markets in the United Kingdom, Urban Education, 39, pp. 537-560.

SAPORITO, S and LAREAU, A (1999) School Selection as a Process: The Multiple Dimensions of Race in Framing Educational Choice. Social Problems, Vol. 46, No. 3 (Aug., 1999), pp. 418-439

SELLSTROM, E. and BREMBERG, S. (2006) Is there a "school effect" on pupil outcomes? A review of multilevel studies, *J Epidemiol. Community Health*, 60, pp. 149-155.

SEPPANEN, P. (2003) Patterns of 'public-school markets' in the Finnish comprehensive school from a comparative perspective, Journal of Education Policy, 18, pp. 513-531.

SKOLVERKET (2006). *Equity trends in the Swedish school system A quantitative analysis of variation in student performance and equity from a time perspective, English summary of report nr 275 (Anders Auer)* Stockholm: Swedish national agency for education.

SKOLVERKET (2010).The SIRIS database, http://siris.skolverket.se

STATISTICS SWEDEN (2007). *Barn, boendesegregation och skolresultat (Children, segregated housing and school results)*. Demografiska rapporter, 2007:2.

SUBRAMANIAN, S. V. (2004) The relevance of multilevel statistical models for identifying causal neighborhood effects, Social Science and Medicine, 58, pp. 1961-1967.

SYKES, B. and MUSTERD, S. (2010) Examining Neighbourhood and School Effects Simultaneously: What Does the Dutch Evidence Show?, *Urban Studies*, Published online before print September 6, 2010, doi: 10.1177/0042098010371393 pp.

TAYLOR, C. (2001) The geography of choice and diversity in the 'new' secondary education market of England, Area, 33, pp. 368-381.

WERFHORST, H. G. V. D. & MIJS, J. J. B. (2010) Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective. Annual Review of Sociology, 36, 407-428.

Appendix

Table 6 All students, empty model

	Observed school			Counterfactual school		
	2000	2003	2006	2000	2003	2006
FIXED	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Intercept	-0,017 (0,009)	-0,038 (0,012)	-0,051 (0,012)	-0,004(0,0008)	0,023(0,008)	0,033(0,008)
RANDOM	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation
Variance components						
Student level	93,47%	85,26%	83,71%	95,39%	94,62%	94,79%
School level	6,53%	14,74%	16,29%	4,61%	5,38%	5,21%
Sum	100%	100%	100%	100,00%	100,00%	100,00%
-2*log likelihood	264982,1	293128,8	332647	272536,4	314928,6	359653,6

Table 7 All students, student level model

	Observed school			Counterfactual school		
	2000	2003	2006	2000	2003	2006
FIXED	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Intercept	-0,842 (0,020)	-0,762 (0,019)	-0,886 (0,019)	-0,845(0,02)	-0,704(0,019)	-0,804(0,018)
Student variables						
Born abroad	-0,082 (0,009)	-0,082 (0,009)	-0,094 (0,008)	-0,078(0,009)	-0,067(0,009)	-0,068(0,009)
Female	0,407 (0,006)	0,360 (0,005)	0,381 (0,005)	0,409(0,006)	0,363(0,006)	0,38(0,006)
Parents' education	0,601 (0,006)	0,556 (0,006)	0,551 (0,005)	0,619(0,006)	0,590(0,006)	0,584(0,006)
Social benefit	-0,603 (0,011)	-0,687 (0,011)	-0,656 (0,011)	-0,608(0,012)	-0,709(0,013)	-0,671(0,013)
Student's disp. Income	0,052 (0,002)	0,044 (0,002)	0,055 (0,002)	0,052(0,002)	0,042(0,002)	0,053(0,002)
RANDOM	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation
Variance components						
Student level	77,51%	71,77%	70,45%	79,40%	81,55%	82,00%
School level	3,41%	12,44%	13,54%	2,82%	3,33%	3,62%
Sum	80,92%	84,21%	84,00%	82,22%	84,88%	85,62%
-2*log likelihood	246 518,20	274 615,90	310771,30	254830,6	298836,4	341913

Table 8 All graduating students, student and school level model

	Observed school			Counterfactual school		
	2000	2003	2006	2000	2003	2006
FIXED	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Intercept	-1,099 (0,176)	-0,293 (0,257)	-1,966 (0,267)	-0,48(0,249)	-0,914(0,231)	-0,466(0,233)
Student variables						
Born abroad	-0,076 (0,009)	-0,082 (0,009)	-0,092 (0,008)	-0,072(0,009)	-0,057(0,009)	-0,065(0,009)
Female	0,407 (0,006)	0,359 (0,005)	0,3800 (0,005)	0,409(0,006)	0,363(0,006)	0,381(0,006)
Parents' education	0,592 (0,006)	0,555 (0,006)	0,549 (0,005)	0,613(0,006)	0,582(0,006)	0,577(0,006)
Social benefit	-0,596 (0,011)	-0,682 (0,011)	-0,649 (0,011)	-0,604(0,012)	-0,70(0,013)	-0,667(0,013)
Student's disp. Income	0,052 (0,002)	0,045 (0,002)	0,055 (0,002)	0,052(0,002)	0,042(0,002)	0,053(0,002)
School variables						
Born abroad	-0,036 (0,056)	0,175 (0,087)	0,308 (0,087)	-0,153(0,071)	-0,029(0,067)	0,094(0,069)
Female	-0,062 (0,085)	0,112 (0,121)	0,446 (0,113)	0,049(0,094)	-0,103(0,090)	-0,111(0,092)
Parents' education	0,477 (0,043)	0,089 (0,071)	0,153 (0,067)	0,553(0,057)	0,731(0,057)	0,753(0,058)
Social benefit	-0,225 (0,106)	-0,936 (0,178)	-1,170 (0,172)	-0,033(0,13)	-0,463(0,135)	-0,354(0,146)
Student's disp. Income	0,013 (0,020)	-0,064 (0,029)	0,100 (0,030)	-0,072(0,029)	-0,003(0,026)	-0,076(0,027)
RANDOM	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation	Percent of variation
Variance components						
Student level	77,51%	71,87%	70,45%	79,40%	81,55%	82,00%
School level	2,81%	11,87%	12,88%	2,45%	2,56%	2,86%
Sum	80,32%	83,73%	83,33%	81,84%	84,12%	84,86%
-2*log likelihood	246 367,40	274 571,10	310680,3	254704,6	298608,9	341725,4