

School choice and increasing performance difference: a counterfactual approach

John Östh

john.osth@kultgeog.uu.se

Uppsala University

Department of Social and
Economic Geography

SE-752 20 Uppsala

www.kultgeog.uu.se

+46 18 471 73 88

Eva Andersson

eva.andersson@humangeo.su.se

Stockholm University

Department of Human
Geography

SE-106 91 Stockholm

www.humangeo.su.se

+46 8 16 20 00

Bo Malmberg

sebo.malmberg@humangeo.su.se

Stockholm University

Department of Human
Geography

SE-106 91 Stockholm

www.humangeo.su.se

+46 8 16 48 53

Abstract

In recent years, structural changes to the school system, including the introduction of independent schools, have increased school choice alternatives in Sweden. Consequently, a large share of today's students attend a school other than the one 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—hypothetically—should not increase the between-school variation in grades. In reality, however, between-school variation in grades has increased in recent years. The aim of this paper is to test whether increasing between-school variance can be explained by changes in residential patterns, or if it must be attributed to structural change. Using a counterfactual approach, the students' variations in grades are compared between observed schools of graduation and hypothetical schools of graduation. The multilevel results indicate that school choice seems to increase between-school variation of grades.

Introduction

During the past two decades the Swedish educational system has experienced rapid structural change. Until 1992 all comprehensive education, with very few exceptions, took place in public schools. Swedish municipalities had a great deal of self-governance, and could 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 at 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 much influence over the catchment area.

Independent schools were first introduced in Sweden in 1992. While all education remained free of charge to 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 one containing several schooling agents spurred competition between schools.

In the early years, the few available independent schools played a minor role in Swedish schooling. According to the National Agency for Education, in the year 2000 3.2% of secondary school graduates attended independent schools. This percentage had increased to 6.3% by 2003 and 8.5% by 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 the 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. Profiling can entail, for example, a focus on specific school subjects (sports, language, etc.), the use of specific pedagogical methods (such as Montessori), or an adherence 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 and sports orientations, for instance. As a result of this increasing competition, parents and students receive advertisements from the different schools (public and independent), and the parents fill in forms for the student's school choice. These choices are typically made when the child is about to start first grade, or is about to leave the third, sixth or ninth grade.

During this period of structural transformation of the Swedish school system, it has been possible to observe a dramatic increase in between-school variance in student performance, as well as a decline in mean student performance; see Figure 1.

Figure 1 approximately here

In reading tests, for example, between-school variance in performance in 2000 was below 10%, compared to a mean of 36% in the OECD sample, but has been rising rapidly since then. Given that Luyten et al. (2005) have demonstrated that countries with high levels of between-school variance tend to do less well in terms of overall student performance (see Figure 2), it is possible that the observed increase in between-school variance has not been good for overall student performance.

Figure 2 approximately here

Therefore, in order to evaluate the effects of the Swedish school reforms it is important to analyse the extent to which increasing variance in school performance can be attributed to structural change in the school system, or whether there are other factors involved. Here, Lindbom (2010, p. 624f) argues that increasing residential segregation is the main reason for increasing ethnic school segregation in Sweden's metropolitan areas, whereas Bunar (2010a; 2010b) claims that school choice has been the main driver of increased school segregation.

This paper will not directly address the driving forces behind increasing ethnic segregation in Swedish schools, but instead the related question of why Swedish schools have become increasingly differentiated in terms of performance. More precisely, we will take our starting point in Lindbom's claim that increasing school differentiation can result from increasing residential segregation. Thus, the aim of this paper is to test whether increasing between-school variance can be explained by changes in residential patterns, or if it must be attributed to structural change in the school system, including liberalized school choice.

To do this, like Allen (2007) and Lindbom (2010) we have chosen a counterfactual approach. Using data for most ninth-grade students in Swedish schools, we estimate between-school variance in performance both for observed schools of graduation and for hypothetical, close-to-home schools to which students have been grouped based on their home address. If, as suggested, residential segregation (as opposed to school choice and other structural changes to the school system) is the driving force behind increasing performance gaps between schools, then increasing between-school variance should be identified in both the observed schools of graduation and the hypothetical close-to-home schools.

Earlier studies

Internationally, there are important contributions to the analysis of school choice and market-based educational systems (Harris 2012, Alegre and Ferrer 2010, Reay 2004, Taylor 2001). There are also geographers in the field of school choice research who argue 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 (Noreisch 2007; Saporito, Lareau 1999) and/or choosing between public and private alternatives (Seppanen 2003; Karsten et al., 2003).

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

Since the 1990s, multilevel analysis has become the standard tool for differentiating school-level and student-level effects on educational outcomes. For example, in the well-cited PISA studies much of the results are based on multilevel analysis, and this is of course also the case for PISA results based on Swedish data. Apart from the PISA studies, there is also a handful of other multilevel studies of school effects in Sweden. Dryler (2000) studies ethnic

segregation between schools and whether or not it has an effect on youths' grades, incomplete grades and the probability of passing on to theoretical programs in upper secondary education. Dryler finds positive effects of the percentage of foreign-born on educational outcomes in Sweden: she finds that large proportions of foreign-born affected average grades negatively in 1990/91. Interestingly, these differences had disappeared in students' average grades in 1997 (compare to 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 at a given school tend to depress grades among the students there in general.

A last example of Swedish school effects also shows the result that individuals and their households did not fully explain the school variation (Statistics Sweden (SCB), 2007). That is, multilevel modelling is motivated and there is reason to explore differences in school allocation under different school regimes (i.e., choice, non-choice). In the SCB study several variables are controlled for, including parental education, family composition and ethnic background, time in Sweden and 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 a 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 detects a substantial increase over the time period 1998-2004 (Skolverket, 2006). Consequently, according to

the Agency, the opportunity for students to learn will not be equal and will depend to a greater extent on which school the students attend. The study is based on students' grades as well as 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?

School and neighbourhood effects

Of particular interest in our comparison of between-school variation in observed and hypothetical close-to-home schools are studies that analyse both neighbourhood- and school-level effects on educational performance. The reason is that the "school effects" we estimate for our hypothetical close-to-home schools can in fact be regarded as the effects of sharing the same neighbourhood. To our knowledge, there are only two Swedish studies comparing school and neighbourhood effects (Bergsten 2010, Brännström, 2008).

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

The contexts on both school and neighbourhood levels were characterized 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 are neighbourhoods. School effects were around 23%, and neighbourhood effects around 3% (compare to empty model in this paper). When controlling for

individual variables the estimate decreased considerably (school effect around 7%) 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%. (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 consider. (We do compare observed school population with a counterfactual close-to-home school population and, as stated earlier, 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 explains about 21% of the total variance in grades among upper secondary school students, whereas the neighbourhood level explains 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 highly correlated to the population in the neighbourhood.

Similar to Brännström (2008) and Bergsten (2010), researchers examining educational opportunities and neighbourhood effects increasingly acknowledge the importance of schools as mediating the effects of neighbourhoods. Some attribute greater importance to the school in shaping students' future achievements. In his study on secondary education in Helsinki, Kauppinen (2008) finds that the effect on a neighbourhood's educational composition on

educational choice is predominantly mediated by the school's socio-economic composition (Kauppinen, 2008).

The finding that schools play a more important role than neighbourhoods in educational performance indicates that we should not expect increasing residential segregation to have played an important role in increasing between-school variance among Swedish schools.

The literature also contains examples of studies that use a counterfactual approach to study the effect of school choice, like we have done (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 more closely at the results found by Bernelius (2005). Bernelius' study from Helsinki provides strong evidence of school choice contributing to increasing differentiation in 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 socioeconomic family status. According to Bernelius, educational performance 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 to build on individual-level data on each student's elementary school (grades 1 to 6), their current secondary school (grades 7 to 9) and their educational outcomes at the beginning of seventh grade.

In a Swedish study, Lindbom and Almgren (2007) investigate the difference between observed and close-to-home schools concerning *student composition* in the school area. Lindbom and Almgren find an increase in ethnic segregation and no significant change in socioeconomic 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 more complete data material on schools instead of areas surrounding schools, and on cities other than the three largest in Sweden.

In her study of the English secondary school system, Allen (2007) finds that around 50% 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 confirm 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 kinds of schools. Allen's study is comparable to ours in that it aims to illustrate how changes to the school system have altered the allocation of students to schools.

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 their 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.

Our study includes all graduating secondary school students in Sweden for the years 2000, 2003 and 2006. 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, students in smaller urban areas and in rural districts have longer distances to travel and fewer schools to choose between. However, the problem is reduced because of two factors. First, a school bussing system shuttles students from smaller urban and rural areas 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 does not risk an overestimation of the effects of school choice on the variation in grades.

To assess the effects of school choice on students' performance each student is 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 is assigned to each student. Schools differ

significantly in size and, consequently, a small school cannot be assigned to all neighbouring students simply because it is located closer to their home than all other school alternatives. To solve the school size constraint a computer program was developed, whereby 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 the basis of proximity. In the first round all students living within 50 metres of a school are assigned to it, then in the second round all students within 100 metres, and thereafter 150 metres, 200 metres and so on. When a school reaches its maximum student capacity, no more students are assigned to that school; any remaining students are assigned to other relatively close-to-home schools that still have vacancies. The school-matching iterations continue until every student has 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 of any municipality is essentially the same as the total number of students. Consequently, assigning students to schools is a swift procedure in which more than 99% of the students are assigned to their closest, second-closest or third-closest-to-home school. All students are assigned to schools to which they would be 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 of the construction of school catchment areas. In the lower part, a close-up of the 50-metre circles is illustrated. In 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 (close-to-home) school of graduation. 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 approximately here...

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). Students' performance is such an example. A student's performance can partly be attributed to his/her own capacity, environment, parents, etc. However, it may also partly be explained by the school context in which he/she is studying. In order to better estimate students' performance, the analysis must encompass the underlying nested hierarchical structure. In this study, two hierarchies are used: the individual level and the school level (see Table 1 for number of students and schools each year). The results of the multilevel analysis will distinguish the extent to which the individual and the school, respectively, affect the students' performance.

The analyses in this paper are conducted in three steps. In the first step, the variation in grades among Swedish students during the three years of study is analysed using no explanatory variables (empty model). This kind of analysis is used not to explain variation in grades (since error, or variation, equals 100%) but rather to assess the extent to which grade variation is determined by variation on school or individual level. If school equity is high and no (or marginal) sorting of students exists, school-level variation is low; but if the reverse situation is true,

school-level variation will make up a greater share of the total variation. In the remaining two steps explanatory variables, first on individual level and thereafter on individual and school levels, are introduced. Variation on school and individual level will drop as a result of the introduction of parameters. Since the same set of variables are used to analyse variation among both the observed and counterfactual populations, the school- and individual-level variation for both populations and all years can be used to assess whether the sorting of students has increased or decreased over time and whether sorting is different in the two datasets.

A multilevel model can be formulated as:

$$y_{ij} = \beta_0 + \beta_{ij}x_{ij} + \beta_jx_j + u_j + e_{ij} \quad (1)$$

Here y_{ij} is the standardized grade of each student, β_0 is the intercept, x_{ij} represents individual level covariates, x_j represents school level covariates, u_j is the school-specific random effect, and e_{ij} is the individual-specific error.

In our paper we first estimate an ‘empty model’ without covariates. The empty model is:

$$y_{ij} = \beta_0 + u_j + e_{ij} \quad (2)$$

The empty model is used to construct a base model of grade variation in Swedish schools by which the variations on school and individual level can be compared.

In the next step we estimate a model with covariates, that is (1).

The estimates are done separately for the case where the school j is the observed school and where j is the counterfactual school.

We run all regressions using two different populations: first, all students graduating from schools with a minimum of 15 graduating students; second, 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. Results of analyses in which all schools/students are included display similar patterns of coefficient strengths and direction and school-level and student-level variations. Results from regressions using all graduating students are available on request .

All data used in this study were collected from the PLACE database (located at the Department of Social and Economic Geography, Uppsala University); see Table 1. The database contains detailed annual information on all Sweden-resident individuals' work-related, socioeconomic and educational status for the years 1990 through 2006. The material in this study makes use of the following variables, presented in Table 1.

Students' graduating scores from secondary school are used to test the variance in grades between students and schools in 2000, 2003 and 2006. The same grading system has been used throughout the years in the study. The grades describe the sum of scores from different subjects. The following subject scores are given: 0 for fail, 10 for pass, 15 for pass with distinction, and 20 for pass with special distinction (which is the official weighting of grades in Sweden). 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. In order to increase the validity of over-time comparisons the grades

are Z-score transformed, assuring that the average grade is the same throughout the years studied.

Measuring variation in grades for observed schools of graduation is relatively straightforward. However, doing so for counterfactual close-to-home schools of graduation is more complex. Students' observed grades are largely a result of the individual student's ability, ambition and socioeconomic background, but also of teachers' skills as well as fellow students' abilities and opinions regarding education. Using observed grades in the counterfactual close-to-home schools means that we implicitly include the average observed school effect. In order to control for this, we make use of a dependent variable (grade) in which the observed school effect is subtracted from each student's grade. The dependent variable used in analyses of counterfactual close-to-home schools' regressions can be expressed as $y_{ij} - u_k$, where u_k is derived from the corresponding observed school multilevel regression¹.

In the *Results* section, the effects of using both kinds of dependent variables for the regression results on counterfactual schools are discussed further.

Beside variables depicting grades and schools, five variables, representing the students' socioeconomic background are also included; see Table 1. The students' socioeconomic status variables include, firstly, the variable *foreign-born*, which distinguishes between Swedish-born individuals and individuals born in any other

To exemplify, the observed school level variation (u_j) in the multilevel model $y_{ij} = \beta_0 + \beta_{ij}x_{ij} + \beta_jx_j + u_j + e_{ij}$ is saved and subtracted from each individual's grade. The new, observed school variation subtracted, grade variable is used in the corresponding counterfactual model, $y_{ij} - u_k = \beta_0 + \beta_{ij}x_{ij} + \beta_jx_j + u_j + e_{ij}$. Please note that the observed school level variation u_j is renamed u_k when used in the counterfactual model.

country. Secondly, the *sex* variable distinguishes between boys and girls. Thirdly, *parents' education* shows whether either of the student's parents has a post-upper secondary education. The fourth variable confirms whether the student's household received *social benefits* during his/her year of graduation from secondary school. Lastly, we include the *student's disposable income* during his/her year of graduation (hundred SEK, log value). This has been calculated by Statistics Sweden using household disposable income and the consumption weight of the student. 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 at each school who are either foreign-born or female, have higher-educated parents, or reside in households dependent on social benefits. The *student's disposable income* variable on school level is the mean disposable income among the students at each school. The school-level variables are calculated separately for each of the two datasets (observed and counterfactual).

Table 1 approximately here ...

Results

Multilevel analyses are conducted on two different datasets, observed and counterfactual, for 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. The two datasets cover graduating compulsory school students from all of Sweden.

The results of the empty-model regressions are illustrated in Table 2, and 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 that motivates the use of multilevel models.

The pattern of the observed dataset displays significantly greater school variance than noted for the counterfactuals, which also becomes more pronounced over time. The strong increase in variation in 2003 coincides with the increase of independent schools in Sweden, and more importantly, the public schools' response to increased competition through specialization and profiling (Skolverket 2010). Hence, the empty model indicates that school choice (demonstrated in the observed dataset), not residential segregation (demonstrated in the counterfactual dataset), is the factor that determines the greatest share of variance.

Moreover, between-school variance is much higher (21%) if only schools in municipalities with high shares of students (>20%) attending independent schools are used in the analysis, but relatively low (4.3%) if only schools in municipalities with low shares attending independent schools (<5%) are included (results available from the authors upon request). This strengthens the view that school choice plays an important role in the increasing differentiation of Swedish schools.

Table 2 approximately here ...

The counterfactual analyses presented in Table 2 make use of a dependent variable (standardized grades) that excludes the average school effects of each pupil's observed school of graduation (see discussion in the *Research design*

section). It might be suggested, however, that when the observed school effect is excluded, large proportions of unmeasured, direct and indirect effects of student characteristics are excluded as well. If these excluded characteristics are strong, the inference may be incorrect. In Table 3, counterfactual dataset results are displayed from empty models using dependent variables that are not corrected for observed school effects. The results indicate that the use of non-corrected grade variables makes school-level variance even lower. This reduction in school-level effect is also evident in all other counterfactual models using the non-corrected dependent variable (the variation attributable to schools drops between 1.5% and 2% - these tables are available on request). This means that by excluding the average school effect of the observed school from the dependent variable in the counterfactual models, we avoid exaggerating the difference in school-level variation between the observed and counterfactual models.

Table 3 approximately here...

Table 4 illustrates estimates from the student-level model, including five fixed student-level variables. The results suggest that through the inclusion of 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 level in the counterfactual dataset explains 3% of the variation in all the years, while the school level 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 to 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 *foreign-born* and residing in a home with *social benefits* are negatively associated with good grades. Of the two, the *social benefits* variable has the greatest negative impact on students' grades. In contrast, being a girl and/or a child of parents with higher education is strongly and significantly associated with higher grades. The *student's disposable income* level is positively associated with higher grades, albeit less strong. In summary, the variance in results between observed schools increases greatly, which is not the case for the counterfactual schools. Secondly, this shift was not notably reduced by the introduction of controls on the student level.

Table 4 approximately here....

Table 5 displays estimates from the student- and school-level model, including five fixed student-level and five fixed school-level variables. The results in Table 5 show that the introduction of fixed school-level variables affects the variance components only to a small extent. This is evident in a comparison of the relative small reduction of $-2 \cdot \log$ likelihood in Table 5 with the corresponding $-2 \cdot \log$ likelihood values in Table 4. The small but noticeable reduction in school-level variance (app. 0.3%-0.6%) is similar in both datasets. In the counterfactual dataset, the variance attributable to the school decreases to less than 3%, 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 for the years 2003 and 2006.

The school-level variables are essentially aggregate student-level variables, depicting the proportion of the various groups of students at each school. School-level variables point at the student's social and economic context. The results indicate that *parents' education* is positively and significantly correlated with higher grades on both the student and the school level in almost all of the observed and counterfactual regressions. Interestingly, parental education shows a strong neighbourhood effect (counterfactual schools). Similarly, receiving *social benefits* negatively correlated with grades on both the student and the school level in observed and counterfactual models. The remaining three variables are less straightforwardly interpreted. The student-level estimates are significant for *foreign-born*, *female*, and *student's disposable income* in both observed and counterfactual models, but the school-level estimates are not (with some exceptions). For the observed schools, all the student-level estimates are significant and have the expected signs. However, school-level effects shift signs or have varying levels of significance. One possible reason for this might be multicollinearity.

Table 5 approximately here...

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% of the total variation is explained. Among other unknown factors, the unexplained variance on student level can be attributable to the student's intellectual ability as well as support from relatives and friends. Due to data limitations, these kinds of variables cannot be included in our analyses. In the counterfactual datasets, almost all variation is observed on the 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 point to the effects of school choice. The choice-driven sorting of students with different abilities and 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 downtown, consequently also means choosing a schooling context that facilitates or impedes one's educational performance.

Concluding discussion

The empirical analysis presented 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. Using a comparison between observed schools and hypothetical, close-to-home schools, we have also been able to show that this trend towards increasing performance gaps cannot be explained by shifting patterns of residential segregation. Between-school variance in student performance is higher for observed schools than for hypothetical schools. This is contrary to what one would find if increasing residential segregation was the critical factor, and the difference also persists when ethnic and socioeconomic background is controlled for. Thus, the most likely driving force behind increasing differentiation is not residential segregation but rather the rapid structural transformation of the Swedish school system. A rapid growth in the number of students attending voucher-financed, independent schools has been an important element in 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 begin reshaping the school system until the early 2000s. That students have increasingly selected non-neighbourhood schools is evidenced by the fact that the median distance from home to school in our data has increased over the period. In 2000 it was 1.6 kilometres, and by 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. But in the Swedish case, this is not what has been observed. With expanding school choice the differences between schools have increased, and at the same time Sweden's comparative performance has declined (OECD, 2010). Thus, as has been the case with other neo-liberal ideas, school choice—when tested—has not been able to deliver the results promised by theoretical speculation.

Moreover, our results demonstrate that an evaluation of trends in between-school variance is of importance when school choice reforms are analysed. Even if there is no dramatic increase in socioeconomic and ethnic segregation (see e.g. Harris, 2012), increasing performance inequality can have potentially negative effects on overall efficiency.

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* 42(11), pp 2674-2686.
- 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-skola-langt-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>. Stockholm, 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.J. (2012) Local indices of segregation with application to social segregation between London's secondary schools, 2003 – 2008/9, *Environment and Planning A*, (pp. -), 2012.
- 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. (2010) School Choice in Sweden: Effects on Student Performance, School Costs, and Segregation, *Scandinavian Journal of Educational Research*, 54, no. 6, pp. 615 - 630.
- 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.

- Luyten, J. W., J. Scheerens, et al. (2005). School factors related to quality and equity. Results from PISA 2000. Paris, OECD.
- NOREISCH, K. (2007) School catchment area evasion: the case of Berlin, *Journal of Education Policy*, 22, pp. 69-90.
- OECD (2010). PISA 2009 results. Vol. 5, Learning trends : changes in student performance since 2000. Paris, OECD.
- 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.

Table 1. 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	Number of students	Count (94,530 in year 2000, 105,899 in 2003, 120,472 in 2006)
	Foreign-born	Swedish-born (0), born in any other country (1)
	Sex	1= female, 0=male
	Parents' education	If either of the student's parents has a post-upper secondary education, the parents' education variable value equals 1. In all other cases, 0.
	Social benefits	If the student's household receives social benefits during the year of graduation from secondary school, the <i>Social benefits</i> variable equals 1. In all other cases, 0.
	Student's disposable income	Calculated by Statistics Sweden based on household disposable income using the consumption weight of the student, hundred SEK, log value
2. School	Number of schools.	Count (1,088 in year 2000, 1,254 in 2003, 1,194 in 2006)
	Foreign-born	Share of foreign-born students at each school
	Sex	Share of female students at each school
	Parents' education	Share of students at each school with parents who have a post-upper secondary education
	Social benefits	Share of students' households receiving social benefits at each of the schools
	Student's disposable income	Mean student disposable income (log of hundreds of SEK) at each of the schools
Dependent variable	Grades	Observed school models: Sum of grades in different subjects (Z-score of grades, annual overall mean = 0)
		Counterfactual school models: Sum of grades in different subjects (Z-score of grades) minus observed school effect.

Table 2. 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

Table 3. Empty model. Counterfactual models make use of a non-corrected dependent variable.

	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)	-0.004 (0.006)	-0.002 (0.006)	-0.001 (0.006)
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%	96.98%	96.75%	96.96%
	0.930 (0.004)	0.891 (0.004)	0.885 (0.004)	0.963 (0.004)	0.953 (0.004)	0.956 (0.004)
School level	6.06%	12.65%	13.49%	3.02%	3.25%	3.04%
	0.060 (0.003)	0.129 (0.006)	0.138 (0.006)	0.030 (0.002)	0.032 (0.002)	0.030 (0.002)
Sum	100%	100%	100%	100%	100%	100%
-2*log likelihood	263304.6	291239.8	330273.4	266022.8	296764.4	338033.1

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						
Foreign-born	-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 benefits	-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

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						
Foreign-born	-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 benefits	-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						
Foreign-born	-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 benefits	-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

Figures

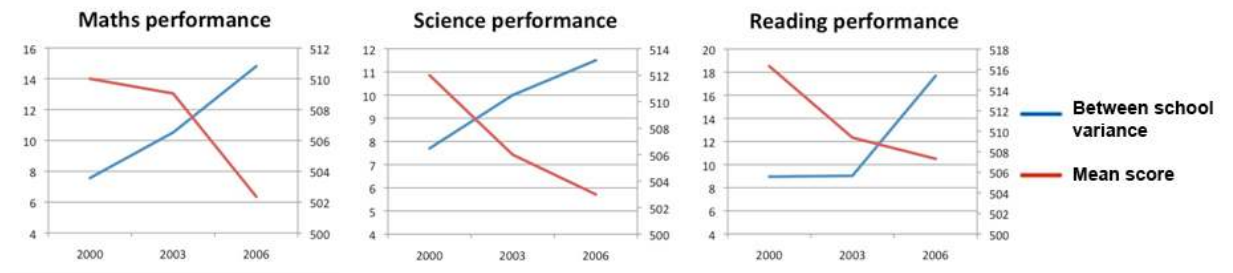


Figure 1. Mean and between-school variance in student performance, Sweden 2000-2006.

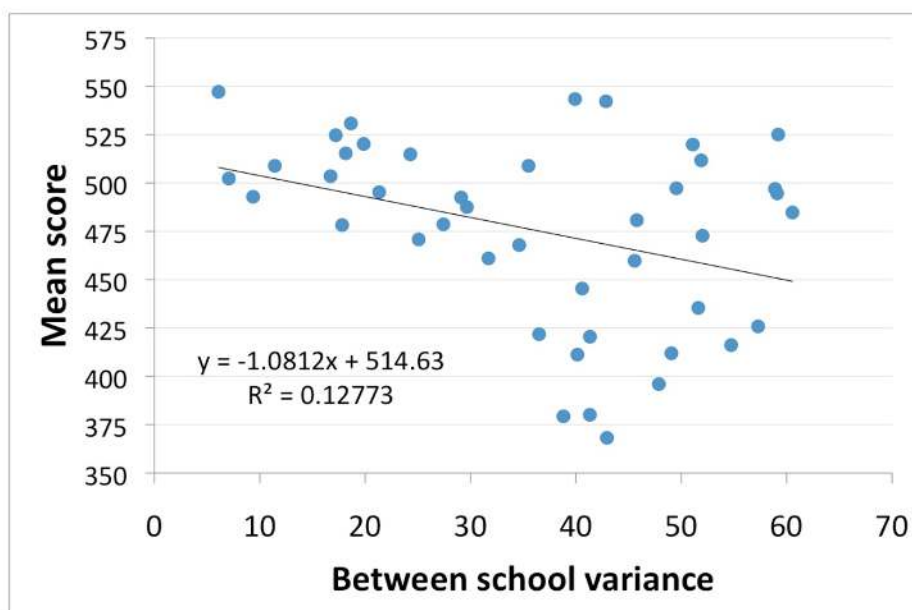


Figure 2. Between-school variance and mean student performance in OECD/Pisa data, 43 countries 2000-06. Robust standard error for the slope parameter is 0.358 (Huber-White).

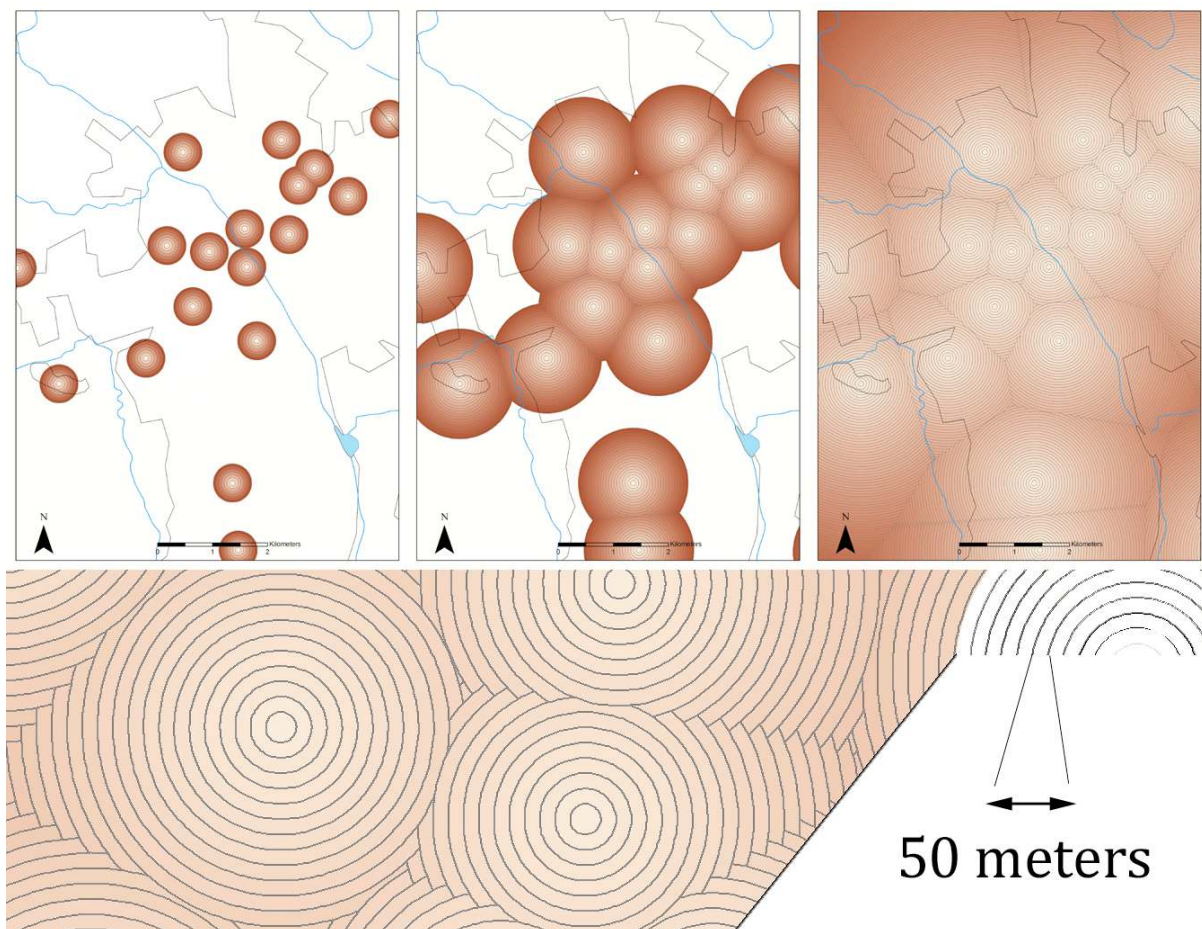


Figure 3. Illustration of the principle behind assignment areas for the creation of close-to-home schools. The central-most part of each circle represents the location of a school. Additional circles represent enlargements in the search areas by 50 meters. As long as a school has not reached its maximum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