

Paweł BUKOWSKI*

 0000-0003-3795-6308

Student Mobility and Sorting of Students¹

Abstract: This study investigates whether improving student mobility leads to greater sorting of students between schools and classes. I isolate an exogenous change in student mobility using the two-stage design of the Polish comprehensive education system and differences in school density across geographic areas. I construct a novel measure of student homogeneity based on Raven's Progressive Matrix test score. One finding is that higher mobility leads to greater sorting of students between schools. Another, more novel, result shows that mobility also leads to higher sorting within a school (across classes). I provide suggestive evidence that demand for peer quality among students motivates school principals to create selective tracks within comprehensive schools.

Keywords: student mobility, sorting, equality of opportunity, Poland

JEL classification codes: I20, I21, I24, I28

Article submitted May 1, 2020, revision received July 17, 2020,
accepted for publication July 21, 2020.

* Centre for Economic Performance, London School of Economics and Political Science, United Kingdom; Institute of Economics, Polish Academy of Sciences, Poland, e-mail: p.bukowski@lse.ac.uk

¹ I thank Sascha O. Becker, David Card, Roman Dolata, Bernhard Enzi, Hedvig Horvát, Gábor Kézdi, Randy Filer, Daniel Munch, Sergey Lychagin, Jesse Rothstein, Christopher Walters, the participants of seminars at Central European University, CERGE-EI and the University of California at Berkeley for their comments and suggestions. I acknowledge the hospitality of the Department of Economics, CLE, IRLE at UC Berkeley. I gratefully acknowledge financial support from CERGE-EI/GDN Grant RRC16+12. All errors are mine.

Mobilność studentów i ich sortowanie

Streszczenie: Artykuł analizuje wpływ mobilności uczniów na ich sortowanie pomiędzy szkoły i klasy. Strategia dla uzyskania związków przyczynowo-skutkowych opiera się na dwustopniowej strukturze polskiego systemu edukacji powszechnej oraz różnicach w gęstości szkół pomiędzy gminami. Miara homogeniczności studentów oparta jest na Matrycach Ravena. Wyniki pokazują, że większa mobilność uczniów zwiększa ich sortowanie pomiędzy szkołami oraz klasami. Dodatkowa analiza sugeruje, że popyt na wysoką jakość rówieśników motywuje dyrektorów szkół do tworzenia specjalistycznych klas.

Słowa kluczowe: mobilność studentów, sortowanie, równość szans, Polska

Kody klasyfikacji JEL: I20, I21, I24, I28

Artykuł złożony 1 maja 2020 r., w wersji poprawionej nadesłany 17 lipca 2020 r.,
zaakceptowany 21 lipca 2020 r.

Introduction

It has been argued that facilitating student mobility could motivate school principals to improve school quality [Friedman, 1955; Hoxby, 2000]. When school funding depends on the number of students and when people care about the quality of education, allowing a free choice of school will drive away some students from their current “low-productive” schools. This process will continue until higher-quality schools dominate the educational market or until schools respond to competitive pressure. Based on this premise, many policies, such as school vouchers or the expansion of school autonomy, have been recently proposed to accelerate student mobility [Gibbons, Machin, Silva, 2008; Kern, Thukral, Ziebarth, 2012].

There is, however, an associated efficiency-equity trade-off as students might sort into schools and classes, affecting equality of opportunity. Advantaged and high-skill children are more likely to select their schools than their disadvantaged and low-skill peers because access to information, performance or ability to commute depend on parental resources [Ajayi, 2012; Ajayi, Friedman, Lucas, 2017; Jensen, 2010]. At the same time, a school has incentives to attract only the best or wealthiest students, so that it performs better on school rankings. Consequently, rich or high-performing students will concentrate in high-quality schools, while poor or low-skilled students will be grouped in low-quality schools, reinforcing their disadvantaged position [Epple, Romano, 1998; Ladd, Fiske, 2000; Hsieh, Urquiola, 2006; Nechyba, 2006; Böhlmark, Holmlund, Lindahl et al., 2015]. Classroom homogeneity might also be affected by the increased mobility. One theoretical argument is that the creation of a high track within a school could be used to attract high-performing or rich students [Epple, Newlon, Romano, 2002]. A second argument is that the creation of homogeneous and easy-to-teach classes might be used to attract high-skilled teachers [Clotfelter, Ladd, Vigdor, 2005].

While numerous studies show that student mobility and school competition lead to sorting of students between schools [Card, Rothstein, 2007; Kalogrides, Loeb, Bêteille, 2013; Collins, Gan, 2013], we know very little about the effect on sorting within a school. The gap in the literature is surprising given that classroom assignment and tracking are of crucial importance for student achievements, as they determine peer composition and teacher quality [Meghir, Palme, 2005; Kremer, Duflo, Dupas, 2011; Figlio, Page, 2002].

This study estimates the effect of student mobility on sorting within and between schools. I employ a difference-in-differences approach based on two institutional features of the Polish education system. First, its comprehensive and obligatory part (grades from 1st to 9th) consists of two separate stages, elementary (grades 1st to 6th) and lower-secondary (7th to 9th). The admission process to the two stages is based on catchment areas with a certain school choice allowed. Second, it is easier for students to use the school choice option at the entrance to lower-secondary education, but only in places with high school density and low transportation costs. The identification strategy is to compare the sorting of students at the entrance to elementary and lower-secondary education in areas with a high availability of alternative schools and low cost of school choice (e.g., urban areas), and juxtapose this difference with the counterfactual difference in areas where school density is low and the cost of commuting is high (e.g., rural areas). The sorting of students is measured using a novel method based on Raven's Progressive Matrix test score. The test captures general intelligence, determined by student genetic abilities and socio-economic background. It is fixed since early childhood, which ensures that the only source of class/school homogeneity is the sorting of students.

The results show that lower-secondary schools are significantly more homogeneous than elementary schools, but only in urban areas, where students can choose from many alternative schools and face a lower cost of school choice. In rural areas, where school choice is scarce and costly, lower-secondary schools are more diverse than elementary schools, a result that can be accounted for by the differences in school sizes. Therefore, in line with the existing literature, I show that improving student mobility leads to higher sorting between schools.

Next, I turn to the analysis of sorting across classes. Conditional on the sorting across schools, classes in urban lower-secondary schools are more homogeneous than in urban elementary schools. Meanwhile, in rural areas, classes in lower-secondary schools are more diverse than in elementary schools. Taken together, these results show that improving student mobility leads to higher sorting within schools. Data on school characteristics is used to test the two theoretical explanations outlined above. The results show that sorting within a school is likely used to attract high-skill or high-income students, which is consistent with a theoretical model developed by Epple et al. [2002]. On the other hand, there is no systematic evidence that school principals attract highly skilled teachers by offering them homogeneous classes.

The paper is organised as follows. The second section depicts the organisation and characteristics of the Polish education system. The third section explains the identification strategy. The fourth section provides the empirical specification and describes the data. The fifth section presents the main results and robustness checks. The sixth section discusses in more detail the effect of student mobility on sorting across classes. Finally, the seventh section concludes.

Institutional Background

The Polish comprehensive education during the analysed period (2010) was compulsory and consisted of six years of elementary education (ISCED 1), followed by three years of lower-secondary education, also called *gimnazjum* (ISCED 2).² Education across the two stages was provided by separate schools, with different managerial and teaching bodies. After finishing the comprehensive part, students could finish education or enrol in academic, mixed or vocational higher-secondary stages (ISCED 3).

The admission processes to elementary and lower-secondary education were the same. Catchment areas were used, meaning that every student from a certain area has a right to attend an assigned local public school. Because there were more elementary than lower-secondary schools,³ the catchment area for the latter was usually larger and contained the catchment areas of several elementary schools. Table 1 shows the ratio of elementary to lower-secondary schools in a rural-urban breakdown and for areas with high and low density of lower-secondary schools. In rural (low density) areas, there are 2.3 (3.1) elementary schools for each lower-secondary school on average and almost 1.5 (1.8) in urban (high density) areas. As an alternative to the local school, parents could request a place in an under-subscribed non-local school, but without guaranteed admission. There were no universal recruitment rules for non-local students. The policy of each school was determined by the principal and a recruitment committee that usually consisted of selected teachers and a school psychologist.

The school principals and the recruitment committee determined classroom assignment. As for lower-secondary education, the most common practice was to create classes with students who are similar in terms of perfor-

² The education system was later reformed in a process that began in 2017 and was scheduled to end in 2020. As part of the reform, the lower-secondary stage was removed and elementary education was expanded from six to eight years. After the 8th grade students now move to a four-year higher-secondary stage. This study analyses data from 2010, which means before the reform.

³ Most of Poland's elementary schools are well established, with a history going back decades. Meanwhile, lower-secondary schools were not established until after 1999. The network of elementary schools thus reflects an outdated demographic model and is relatively dense. The network of lower-secondary schools, in turn, is more "rational" in the sense that it is better adjusted to current demographic needs. Also, elementary education serves younger children for whom the distance to a school matters more than for older children.

mance, foreign language proficiency and place of residence [Szmigel, 2013]. In elementary education, principals could not sort students based on their performance (it was unknown), but they could take into consideration gender composition, place of living and date of birth. Parents had a right to suggest an alternative class assignment. Importantly, the assignment was fixed across grades and subjects, with reallocations allowed only in exceptional cases. The peer composition of classes was thus relatively constant at each stage of education. There were no limits on classroom size.⁴

Table 1. Descriptive Statistics for Municipalities

Variable	Rural (1)	Urban (2)	Low LS/km ² (3)	High LS/km ² (4)
	<i>Total Numbers</i>			
Elementary Schools	10 894	2248	4759	8321
Lower-Secondary Schools	5371	2071	1634	5808
	<i>Averages for Municipalities</i>			
Elementary per Lower-Sec.	2.31	1.49	3.12	1.81
Elementary School per km ²	0.05	0.25	0.03	0.09
Lower-Sec. School per km ²	0.03	0.24	0.01	0.08
Children per Elementary School	153	337	129	191
Children per Lower-Sec. School	186	207	213	161
Public Transportation per km ²	0.07	1.67	0.002	0.26
Tertiary Education Share	4%	11%	3%	5%
Population Density	166	1676	49	399
Population	10 067	156 004	6983	24 096
Number of Municipalities	2386	93	1234	1234

Source: Central Statistical Office of Poland and Herczyński, Sobotka [2013].

Note: Columns (1) and (2) show the descriptive statistics for rural and urban municipalities in Poland, where urban municipalities are those with a population larger than 50,000. Columns (3) and (4) are for areas with a density of lower-secondary schools per km² below and above its median. All the figures are for 2010, except Tertiary Education Share (2002) and Public Transportation km²(2007).

Students were evaluated by two standardised, externally graded and obligatory examinations. After elementary education (6th grade) they took a low-stake exam that served mostly an informational purpose. After lower-secondary education (9th grade) students were tested with a high-stakes exam, which was used in the next stage of education. Based on these two tests, the Ministry of Education estimated the educational value-added measures of the performance of lower secondary schools, which were publicised. School funding

⁴ In 2015, the rules for elementary education were unified and are now based on the date of birth, with an option for parents to request an alternative assignment. As of 2013, a class in grades one to three can have a maximum of 25 students.

was not, however, linked to school performance. In addition, there were various unofficial rankings seeking to assess the average *levels* of elementary or lower-secondary school performance.

There were clear economies of scale for school principals. The central government financed all Polish public schools through a subsidy. In theory, this amount was expected to be sufficient to cover all expenditures on education, excluding investment and pre-school education. In practice, however, it covered only around 50% to 70% of the costs [Herbst, Herczyński, Levitas, 2009; Instytut Badań Edukacyjnych, 2011], while the rest was covered by local governments. Since the governmental subsidy was tied to the student, school funds depended on enrolment. In addition, principals in larger schools had more bargaining power when securing additional funds from the local government. In general, public schools did not advertise themselves, but they could use other ways of signalling their quality. One strategy was “cream-skimming” of students, which improved the position of a school in rankings based on levels. This paper argues that sorting across classes could be one way of attracting high-performing or rich students.⁵

Local governments determined teacher salaries and employment conditions in compliance with a universal collective bargaining agreement called *Karta Nauczyciela*. It specifies the minimum level of wage for each teacher’s rank.⁶ Also, teachers were eligible for overtime pay, monetary awards and other non-monetary benefits, for instance accommodation in a school-owned apartment. Interestingly, despite the fact that prices in rural areas are lower than in cities, teachers working in rural schools receive an extra monetary allowance.

Identification Strategy

Places with higher student mobility may have more homogeneous schools and classes because of other parallel social processes. For instance, similar people tend to live together because of neighbourhood characteristics, local economic conditions or housing prices [Tiebout, 1956]. The quality of local schools influences these characteristics, which might further reinforce self-selection [Figlio, Lucas, 2004; Kane, Riegg, Staiger, 2006]. Consequently, the effect of mobility on sorting will be biased if schools use catchment areas, and places with more mobile students also have more residential sorting.

In order to identify the effect of student mobility on sorting, I exploit the two-stage design of the Polish comprehensive education system and differences

⁵ On the other hand, mixing students across classes might be preferred by “egalitarian” school principals or policymakers as it improves educational equality of opportunity.

⁶ In 2015, the minimum monthly gross wages ranged from PLN 1,513 (EUR 340) to PLN 3,109 (EUR 700). Additionally, the average *total* gross salary for each rank of teacher within a municipality had to be at least as large as specified in the *Karta Nauczyciela*. In 2015, these averages ranged from PLN 2,717 (EUR 612) to PLN 5,000 (EUR 1,126).

in the cost of school choice across rural and urban areas. Since the elementary and lower-secondary education levels are obligatory and locally provided, the homogeneity of classes or schools at both these stages of education is equally influenced by fixed area characteristics. The first assumption is that the difference between the homogeneity of classes or schools across these stages of education is an outcome of changes in student mobility, changes in catchment areas, and changes in classroom/school assignment. The second assumption is that in areas with high costs of school choice (rural areas), the difference in student mobility will be irrelevant. The third assumption, which I later relax, is that changes in catchment areas and classroom/school assignment are the same across different areas.

Consequently, to capture the effect of student mobility on sorting, it is sufficient to compare how sorting differs across stages of education and across areas with different costs of school choice. The identification strategy is summarised in Figures 1 and 2. Each cell lists the forces driving classroom (Figure 1) and school (Figure 2) homogeneity across different stages of education and locations. Such research design can be interpreted as an example of the difference-in-differences technique. The “Treatment” is a change in student mobility. The “Treatment group” is a low-cost area in terms of school choice. “Before and after” are the first and second stages of Polish comprehensive education respectively.

Figure 1. The Identification Strategy – Sorting Within a School

		Higher Student Mobility	
		1 st grade: Elementary	7 th grade: Lower-Secondary
Urban Areas	Residential Sorting	Residential Sorting	Residential Sorting + Classroom Assignment + Ability Sorting across Classrooms
Rural Areas	Residential' Sorting	Residential' Sorting	Residential' Sorting + Classroom Assignment

Source: author’s own elaboration.

Figure 2. The Identification Strategy – Sorting Between School

		Higher Student Mobility	
		1 st grade: Elementary	7 th grade: Lower-Secondary
Urban Areas	Residential Sorting	Residential Sorting	Residential Sorting + Catchment Areas + Ability Sorting across Schools
Rural Areas	Residential' Sorting	Residential' Sorting	Residential' Sorting + Catchment Areas

Source: author’s own elaboration.

In the remaining part of this section, I review key assumptions required for the proposed methodology to provide causal estimates:

Assumption 1. *Treatment* – *students entering lower-secondary education (7th grade) are more likely to use school choice than students entering elementary education (1st grade).*

This assumption is motivated by three arguments. First, students entering lower secondary education are older, which means that commuting is more feasible for them and they are more independent in their decisions about school. Second, their performance is known, which is not the case with younger students entering elementary education. Fewer informational constraints might motivate students to select a better fitting non-local school and allow school principals to screen applicants based on their performance. Third, the catchment area of a lower-secondary school usually contains the catchment areas of several local elementary schools. Consequently, students entering lower-secondary education face larger catchment areas and the composition of their local school will to a lesser extent reflect the residential composition of their neighbourhood.

Table 2. Share of Students in Non-Local Schools

Stage	All	Urban	Rural	High LS/km ²	Low LS/km ²
Elem. School	18%	23%	15%	19%	16%
Lower-Sec. School	24%	42%	16%	31%	17%
Difference	6 pp***	19 pp***	lpp	12 pp**	lpp
N	4907	1524	3383	2540	2367

Source: author's calculation based on the EVA survey.

Note: Columns (Urban) and (Rural) show the statistics for rural and urban schools, where urban schools are those in municipalities with a population larger than 50,000. Columns (Low LS/Km²) and (High LS/Km²) are for areas with a density of lower secondary schools per km² below and above its median. All the figures are for 2010. *** denotes significance at the 1% level, ** at the 5% level.

The higher mobility of secondary-school students estimated from the EVA survey (described in the next section) is documented in Table 2. It reports the proportion of students from the first grade of lower-secondary education attending a non-local lower-secondary school, and also shows what percentage of them attended a non-local elementary school. In the whole sample, 18% of the students went to a non-local elementary school, and 24% went to a non-local lower-secondary school. The difference is highly significant. Moreover, in the next section, Table 4 Column (4), I provide suggestive evidence that the parents of students entering the second stage might be facing fewer informational constraints.⁷

⁷ On the other hand, as reported in Table 1, there are more elementary schools than *gimnazja*, which would make competition among them more likely.

Assumption 2. Control group – *the difference in student mobility across educational stages is irrelevant in areas with a high cost of school choice.*

In certain areas, the school variety is limited and the cost of attending a non-local school is high [Dolata, 2008] because of larger distances and a sparser transportation network. Consequently, even though students entering lower-secondary education are older and have fewer information constraints, they are less likely to exercise their right to request an alternative school. In the empirical part, I focus on two definitions of areas with a high cost of school choice: 1) rural areas and 2) areas with the number of lower-secondary schools per km² below the median. Using data published by Poland's Central Statistical Office, Table 1 provides descriptive statistics for these areas. The rural municipalities,⁸ compared to the urban municipalities, have a three (10) times sparser network of elementary schools (lower-secondary schools), 23 times sparser network of public transportation, and 10 times smaller population density. Similar differences, but of somehow smaller magnitudes, can be found in a comparison of areas with the density of secondaries below the median and those above the median. In line with these arguments, Table 2 reports that in rural areas and areas with low school density, there is no difference in the share of students attending non-local schools between the studied stages of education.

Assumption 3.a. Common trend (for sorting within a school) – *in the absence of an increase in student mobility, a change in class assignment between elementary and lower-secondary education is the same across areas.*

Assumption 3.b. Common trend (for sorting between schools) – *a change in the size of catchment areas between elementary and lower-secondary education leads to the same level of between-school student mixing.*

Assumption 3.a. says that the reasons to sort or mix students across classes regardless of student mobility should be similar in areas with different costs of school choice. The qualitative evidence discussed in Section 6.2 supports this view. Assumption 3.b is similar but considers sorting or mixing across schools. This assumption, however, is not likely to be satisfied. For instance, student mixing should be more intensive in rural areas as there are more elementary schools for each lower-secondary schools there than in urban areas (see Table 1). In other words, the inter-stage difference in school catchment areas will automatically lead to student mixing or sorting. In Section 5.2, however, I account for this problem by assuming that the mixing effect is proportional to the ratio of elementary to lower-secondary schools.

⁸ Those with a population of less than 50,000.

Estimation and Data

The first part of this section explains the measurement of a change in the between-schools sorting of students across stages of education. The second part develops an analogous measure of within-school sorting. The third part presents the data.

Sorting Between Schools

Consider a measure of socio-economic background (SEB) y_{ics} of student i from class c and school s . It can be decomposed into the population mean μ , the school-level deviation from that mean u_s , the class-level deviation from the school mean u_c and the residual component e_{ics} :

$$y_{ics} = \mu + u_s + u_c + e_{ics}$$

By construction, its variance at stage t (either lower-secondary – l or elementary – e) is the sum of the variance of the school-level component, the variance of the class-level component and the residual variance:

$$Var_t = Var_{s,t} + Var_{c,t} + Var_{e,t}$$

For a given educational stage, the intensity of sorting between schools can be defined as the ratio of the school-level variance to the total variance $Var_{s,t}/Var_t$. The change in sorting across educational stages is:

$$\Delta Var_s = \frac{Var_{s,l}}{Var_l} - \frac{Var_{s,e}}{Var_e}$$

Sorting Within a School

The measure of a change in sorting within a school must account for the differences in catchment areas between elementary and lower-secondary education. The intensity of sorting within a school is defined as the ratio of the class-level variance to the total variance $Var_{c,t}/Var_t$. Ignoring the catchment area problem, the change between educational stages is simply $Var_{c,l}/Var_l - Var_{c,e}/Var_e$.

The problem arises because the catchment areas are larger for lower-secondary schools than for elementary schools. When there are no changes in the class composition at the transition between stages, the fraction of variance explained by the school level drops, while the fraction explained by the class level increases correspondingly. To see this, suppose that there is just one class per elementary school and students do not change classmates as they transition from elementary and lower-secondary education. Because the catchment areas of elementary schools are nested within the catchment area of one low-

er-secondary school, the students from multiple elementary schools will go to one lower-secondary school and each class in the latter will consist of students coming from the same elementary school. Consequently, the importance of class level ($Var_{c,l}/Var_c$) increases even though there is no change in student sorting across classrooms.⁹ To correct for this problem, one can adjust for the negative change in the fraction of the variance explained by the schools. I propose the following measure of changes in sorting within a school:

$$\Delta Var_c = \frac{Var_{c,l}}{Var_l} - \frac{Var_{c,e}}{Var_e} + 1[\Delta Var_s < 0]\Delta Var_s$$

where $1_{[a]}$ is an indicator function, taking value zero if expression a is not true and one if true – that is, a change in the fraction of variance explained by school level is negative. Intuitively, the aforementioned problem arises only when lower-secondary schools have larger catchment areas than elementary schools and their ratio $Var_{s,l}/Var_s$ is lower. When there is no change in the class composition, but the catchment areas are larger for lower-secondary schools, $Var_{c,l}/Var_l - Var_{c,e}/Var_e = -\Delta Var_s$ and thus ΔVar_s is subtracted in order to obtain the value of zero. If the catchment areas are the same or sorting across schools overbalances their effect, a simple difference between the fraction of the variance explained by class level captures the effect of interest.

To estimate the proportions of the variance explained by class and school levels, I use a multilevel mixed-effects linear regression (also called a hierarchical linear model). As discussed in the previous section, comparing changes in sorting across areas with different costs of school choice isolates the effect of student mobility.

Data

The data are drawn from a survey conducted by the Educational Value Added (EVA) Team of the Warsaw-based Educational Research Institute. The EVA survey is a random cross-section of Polish students from 2010 and consists of 5,600 first-graders and 5,567 seventh-graders (i.e., the entry grades of elementary and lower-secondary education) from 330 randomly drawn public schools in Poland.¹⁰ The data include a battery of variables on students, parents, teachers, schools and municipalities, including questions about schools' sorting practices. The survey is representative for the Polish population and

⁹ One way of looking at this problem is to realise that, in this scenario, schools in elementary education become classes in lower-secondary education. With one class per elementary school, there is no difference between the terms “school” and “class”. Although there is no change in class composition at the transition to lower-secondary education, the distinction between “school” and “class” begins to matter. This is because groups of students that were “classes/schools” at the elementary stage become “classes” at the secondary stage.

¹⁰ The target population was elementary public schools with first grades larger than 10 students and public lower-secondary schools with seventh grades larger than 20 students.

all the statistics reported in the paper are calculated using the survey weighting scheme. Table 3 summarizes the sample.

The main outcome variable and a measure of the students' socio-economic characteristics is a standardised (separately for the first and seventh graders) cumulative score from Raven's Progressive Matrix test.¹¹ It is designed to capture two abilities: "(a) eductive ability [...] – the ability to make meaning out of confusion, the ability to generate high-level, usually nonverbal, schemata which make it easy to handle complexity; and (b) reproductive ability – the ability to absorb, recall, and reproduce information that has been made explicit and communicated from one person to another" [Raven, 2000: 2]. In other words, eductive and reproductive abilities make it possible to understand concepts and learn new material and are components of an underlying general mental ability (Jensen, 1998). The test usually consists of 4×4 , 3×3 or 2×2 matrix of figures at each entry except the lowest diagonal, which is empty. Figures in each row follow the same pattern, and the task is to identify it and find the missing element. Importantly, Raven's test score is determined only by genetic, parental and environmental conditions during early childhood (Brouwers, Van de Vijver and Van Hemert, 2009). Any post-kindergarten determinants of education, such as school inputs, teacher quality, parental investments or peer effects, should be irrelevant. Consequently, the only reason students might have a similar level of Raven's score is sorting or self-selection. The advantage of Raven's score is that it captures socio-economic characteristics, such as genotype, which are not necessarily captured by other used measures (e.g., mother's education).

In order to test whether Raven's score is not affected by education, Table 4 Columns (1) and (2) regress the mother's or father's education on Raven's score, a dummy denoting observation from lower-secondary education and an interaction term between the two. If education (and other inputs) during elementary lower-secondary education do not matter for Raven's score, there should be no difference in the correlation between parental education and Raven's score between observations from the two stages. Indeed, while there is a naturally positive correlation between the mother's/father's education and Raven's score, it is not significantly different across the stages.

On the other hand, Column (3) shows that the correlation between Raven's score and the desired level of education for a child is larger among lower-secondary than elementary students. The positive coefficient is consistent with the reduced informational constraints faced by parents at the entrance to the former. Since, as reported in Column (4), there is a positive correlation between the sixth grade GPA and Raven's score, students with a higher Raven's score are on average performing better, and their parents might desire a higher level of education for them. Student performance is unknown at the entrance

¹¹ For each student, I calculate the z-score of Raven's test.

to elementary education, and so the correlation between Raven’s score and the desired level of education is significantly lower.

Results

The first part of this section presents the decomposition of the variance of Raven’s score and translates it into the effect of student mobility on the sorting of students. The second part presents the robustness checks.

Decomposition of the Variance of Raven’s Score

Table 5 presents the proportion of the variance of Raven’s score explained by the school and class levels, broken down by the stages of education, and by urban and rural areas. Table 6 presents similar estimates for areas with above- and below-median levels of lower-secondary school density. The proportions and standard errors are estimated using the mixed effect model, weighted by the survey weights. Figures 3, 4, 5 and 6 visualize the results.

Table 3. Descriptive Statistics of the Sample

Variable	Elementary School					Lower-Secondary Schools				
	Obs.	Mean	St. Dev.	Min	Max	Obs.	Mean	St. Dev.	Min	Max
<i>Full sample</i>										
Raw Raven’s score	5589	27.42	8.38	1	59	4907	45.27	7.58	9	60
Respondents per school	5749	36.17	10.04	8	56	4916	34.39	7.14	10	58
Respondents per class	5749	19.35	4.17	8	30	4916	17.81	4.09	6	30
Numbef of schools	180					150				
<i>Urban</i>										
Raw Raven’s score	2103	29.16	8.31	9	55	1524	46.32	7.48	9	60
Respondents per school	2181	39.83	8.25	10	56	1526	35.48	8.53	10	58
Respondents per class	2181	20.2	4.23	8	28	1526	18.26	4.71	8	30
Numbef of schools	58					46				
<i>Rural</i>										
Raw Raven’s score	3486	26.38	8.24	1	59	3383	44.79	7.57	10	60
Respondents per school	3568	33.94	10.37	8	50	3390	33.9	6.36	15	49
Respondents per class	3568	18.84	4.06	8	30	3390	17.6	3.75	6	28
Numbef of schools	122					104				
<i>Low LS/km²</i>										
Raw Raven’s score	2851	26.28	8.18	1	51	2367	44.29	7.82	10	60
Respondents per school	2851	33.61	10.29	8	47	2367	33.77	6.45	16	49
Respondents per class	2851	18.56	4.09	8	29	2367	17.57	3.89	0	28
Numbef of schools	99					73				

cont. table 3

Variable	Elementary School					Lower-Secondary Schools				
	Obs.	Mean	St. Dev.	Min	Max	Obs.	Mean	St. Dev.	Min	Max
<i>High LS/km²</i>										
Raw Raven's score	2738	28.62	8.42	9	59	2540	46.17	7.23	9	60
Respondents per school	2738	36.86	9.05	8	54	2540	34.84	7.65	10	58
Respondents per class	2738	19.19	4.21	8	28	2540	17.97	4.26	8	30
Number of schools	81					77				

Note: Urban (rural) schools are in municipalities with a population larger (smaller) than 50,000. High (low) LS/km² schools are in municipalities with a density of lower-secondary schools per km² above (below) its median. All the figures are for 2010. Unweighted statistics.

Source: author's own calculations.

Table 4. Raven's Score and Education

<i>Dependent Variable:</i>	Mother's Education (1)	Father's Education (2)	Desired Education for a Child (3)	6th grade GPA (4)
Raven's Score	.557 (.042)***	.543 (.040)***	.464 (.035)***	.532 (.017)***
Lower-Secondary	-.265 (.072)***	-.219 (.072)***	-.352 (.065)***	
Raven's Score × Lower-Sec.	-.019 (.051)	-.008 (.050)	.370 (.042)***	
N	10320	10167	10376	4896
Estimator	O Logit	O Logit	O Logit	OLS

Notes: The table shows regressions of the dependent variables on the standardized Raven's Progressive Matrix Test score, a dummy indicating observation from the seventh grade – lower-secondary school (the excluded category is the first grade – elementary school), and the interaction between them. The Mother's and Father's Education are categorical variables, which take values between 1 and 9, where 1 is unfinished elementary education and 9 is PhD. The Desired Education for a Child is a categorical variable, which takes values between 1 and 7, where 1 is vocational education and 7 is PhD. The 6th grade GPA is the average of grades from various subjects, it ranges between 2 and 6, where 2 is the worst score. Robust and corrected for the survey design standard errors are reported in the parentheses. In columns (1) to (3) the numbers show the coefficients from the Ordered Logit regression. *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.

Source: author's own calculations.

In urban areas, at the entrance to elementary education, the school level explains 13%, and the class level explains 1% of Raven's score variation. At the entrance to lower-secondary education, the proportions increase to 28% and 9% respectively. This means that lower secondary schools and classes are more homogeneous than in the case of elementary education. Consequently, the explained proportion of the variance grows from 14% to 37%. The same pattern, but smaller in magnitude, is documented for areas with high school

density. The explained proportion increases from 19% to 23%, even though the fraction explained by the school level drops from 19% to 17%.

The increase in homogeneity is due to increased student mobility and other grade variant changes in school and class assignment (Assumption 1). To isolate the former mechanism, I compare this difference to the difference for the control areas, which did not experience a change in student mobility, that is, for rural areas or areas with below-median lower-secondary school density (Assumptions 2, 3.a and 3.b). At the entrance to elementary education in rural areas, the school and class levels explain 26% and 1% of Raven’s score variation respectively, and in areas with low school density they explain 25% and 2% respectively. At the entrance to lower-secondary education, the importance of the school level drops to 5% in both areas, which means that lower-secondary schools are more heterogeneous than elementary schools. The drop is likely a result of the differences in the sizes of catchment area across the stages of education, which I quantify in the next section.

Table 5. Decomposition of the Variance of Raven’s Score – Urban vs. Rural

<i>Dependent Variable:</i>	Proportion of Variance Explained (1)	Robust St. Errors (2)	95% C.I. Lower Bound (3)	95% C.I. Upper Bound (4)
<i>Elementary School – Urban</i>				
School level $Var_{s,e}/Var_e$.1258	.0268	.0828	.191
Class level $Var_{c,e}/Var_e$.0145	.0112	.0032	.0659
Residual	.8598	.0257	.8108	.9117
<i>Lower-Secondary School – Urban</i>				
School level $Var_{s,l}/Var_l$.2768	.1011	.1353	.5663
Class level $Var_{c,l}/Var_l$.0936	.0294	.0505	.1733
Residual	.6297	.0502	.5386	.7362
<i>Elementary School – Rural</i>				
School level $Var_{s,e}/Var_e$.2581	.0461	.1818	.3664
Class level $Var_{c,e}/Var_e$.0135	.0079	.0043	.0423
Residual	.7284	.0298	.6722	.7893
<i>Lower-Secondary School – Rural</i>				
School level $Var_{s,l}/Var_l$.0535	.0142	.0318	.0899
Class level $Var_{c,l}/Var_l$.06	.0156	.0361	.0997
Residual	.8865	.0333	.8236	.9543

Notes: The table shows the decomposition of the variance of the standardized Raven’s Progressive Matrix Test Score by the school and class level. The estimation was conducted using the mixed (hierarchical) effect model. Each stage was weighted using the survey weighting scheme. Urban (rural) schools are in municipalities with population larger (smaller) than 50,000.

Source: author’s own calculations.

The increase in homogeneity is due to increased student mobility and other grade variant changes in school and class assignment (Assumption 1). To isolate the former mechanism, I compare this difference to the difference for the control areas, which did not experience a change in student mobility, that is, for rural areas or areas with below-median lower-secondary school density (Assumptions 2, 3.a and 3.b). At the entrance to elementary education in rural areas, the school and class levels explain 26% and 1% of Raven's score variation respectively, and in areas with low school density they explain 25% and 2% respectively. At the entrance to lower-secondary education, the importance of the school level drops to 5% in both areas, which means that lower-secondary schools are more heterogeneous than elementary schools. The drop is likely a result of the differences in the sizes of catchment areas across the stages of education. I quantify the influence of these differences in the next section.

Table 6. Decomposition of the Variance of Raven's Score – High vs. Low School Density

<i>Dependent Variable:</i>	Proportion of Variance Explained (1)	Robust St. Errors (2)	95% C.I. Lower Bound (3)	95% C.I. Upper Bound (4)
<i>Elementary School – High LS/km²</i>				
School level $Var_{s,e}/Var_e$.1855	.0451	.1152	.2987
Class level $Var_{c,e}/Var_e$.0081	.0084	.0011	.0612
Residual	.8063	.0345	.7415	.8768
<i>Lower-Secondary School – High LS/km²</i>				
School level $Var_{s,l}/Var_l$.1675	.0581	.0849	.3307
Class level $Var_{c,l}/Var_l$.065	.0189	.0367	.115
Residual	.7675	.0409	.6915	.8519
<i>Elementary School – Low LS/km²</i>				
School level $Var_{s,e}/Var_e$.2452	.052	.1619	.3715
Class level $Var_{c,e}/Var_e$.0215	.0101	.0086	.0542
Residual	.7332	.0277	.6809	.7895
<i>Lower-Secondary School – Low LS/km²</i>				
School level $Var_{s,l}/Var_l$.0488	.0167	.0249	.0955
Class level $Var_{c,l}/Var_l$.0788	.0212	.0466	.1334
Residual	.8724	.039	.7993	.9522

Notes: The table shows the decomposition of the variance of the standardized Raven's Progressive Matrix Test Score by the school and class level. The estimation was conducted using the mixed (hierarchical) effect model. Each stage was weighted using the survey weighting scheme. High (low) LS/km² schools are in municipalities with a density of lower-secondary schools per km² above (below) its median.

Source: author's own calculations.

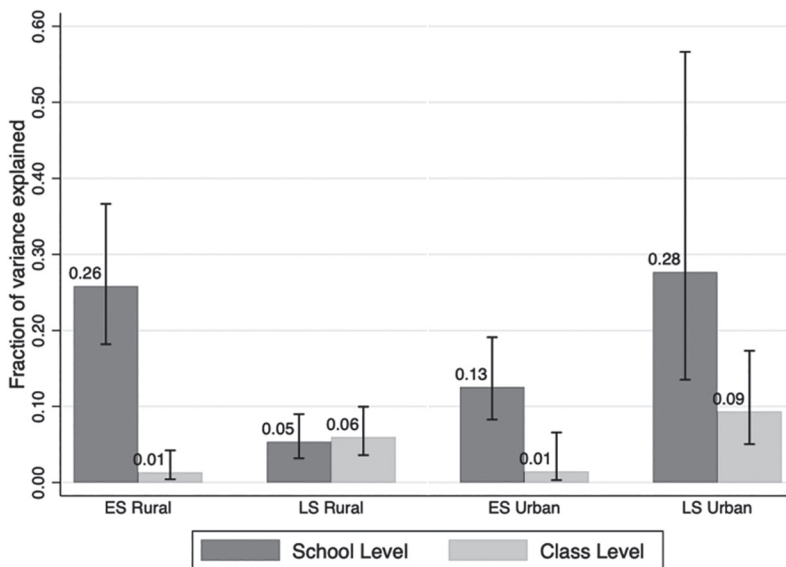
The fraction explained by the class level rises to 6% in rural schools and 8% in areas with low school density. The interpretation of this change, as dis-

cussed in Section 3, is less straightforward. Suppose that there is just one class per elementary school and students do not change classmates as they transition from elementary and lower secondary education. Because the catchment areas of elementary schools are nested within the catchment area of one lower-secondary school, students from multiple elementary schools will go to one lower-secondary school and each class in the latter will consist of students coming from the same elementary school. Consequently, the importance of the class level increases even though there was no change in class composition. However, this also implies that the unexplained part of the variance does not alter. Contrary to this, Figures 4 and 6 document an increase in the unexplained part of the variance, which means that classes are more heterogeneous at the entrance to lower-secondary education compared with elementary education. Based on Equation 4, the drop in sorting within a school is 16 *pp* for rural areas and 14 *pp* for low school density areas.

Student Mobility and Sorting of Students

This section interprets the above results in light of the identification strategy outlined in Section 3. Ignoring the control group,¹² the difference between the stages of education for areas with a low cost of school choice is interpreted as the lower bound of the potential effect of mobility on sorting.

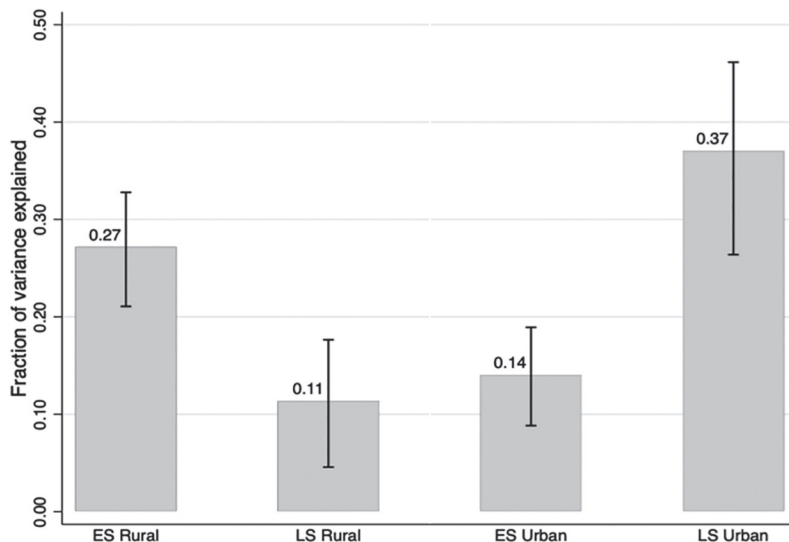
Figure 3. Decomposition of the Variance of Raven’s Score – Urban vs. Rural



Source: author’s own calculations.

¹² That is, only Assumption 1 is satisfied.

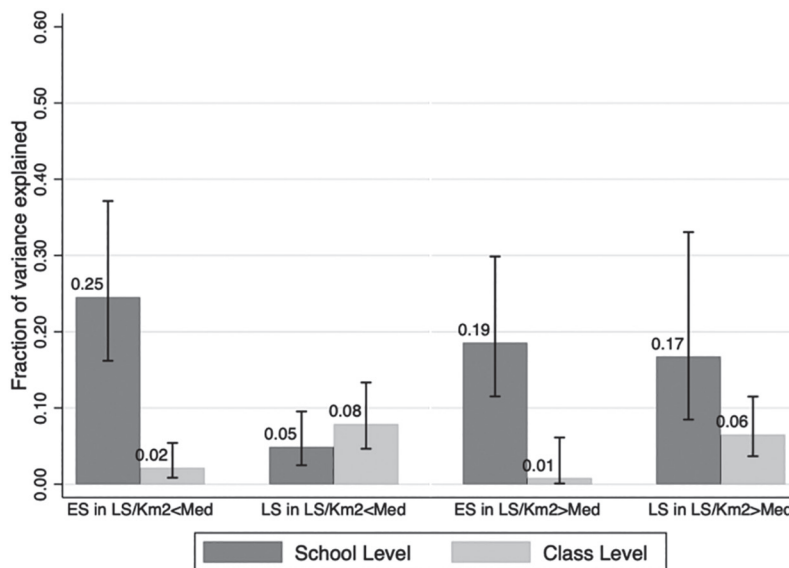
Figure 4. Residual fraction of the Variance of Raven, Score – Urban vs. Rural



Note: Figures present the decomposition of the variance of the standardized Raven’s Progressive Matrix Score, by school and class level, and the unexplained (residual) proportions. The estimation uses the mixed (hierarchical) effect model. The urban (rural) schools are in municipalities with population larger (smaller) than 50,000.

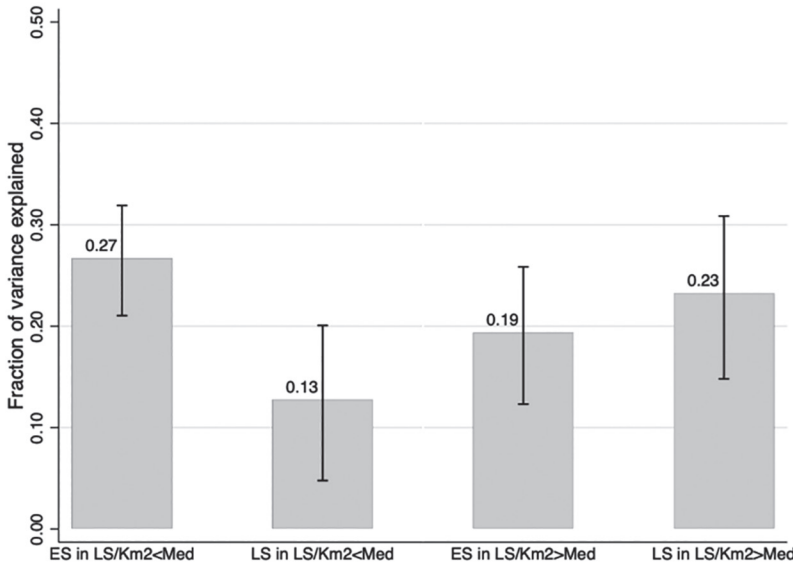
Source: author’s own calculations.

Figure 5. Decomposition of the Variance of Roven’s Score – High vs. Low School Density



Source: author’s own calculations.

Figure 6. Residual fraction of the Variance of Raven’s Score – High vs. Low School Density



Note: Figures present the decomposition of the variance of the standardized Raven’s Progressive Matrix Score, by school and class level, and the unexplained (residual) proportions. The estimation uses the mixed (hierarchical) effect model. The high (low) LS/km² schools are in municipalities with the density of lower-secondary schools per km² above (below) its median.

Source: author’s own calculations.

The reason is that it ignores the mixing effect of catchment areas and general attempts to equalise classes in lower-secondary education. When comparing urban and rural areas, the importance of the school level (ΔVar_s) increases by 15 *pp* and the importance of class level grows (ΔVar_c) by 8 *pp*. Comparing this with the counter-factual difference in the control group¹³ provides an upper bound estimate. For the urban vs. rural comparison, the change in sorting between is $\Delta Var_s^{URBAN} - \Delta Var_s^{RURAL} = 15 \text{ pp} - (-21 \text{ pp}) = 36 \text{ pp}$, whereas the change in sorting within is $\Delta Var_c^{URBAN} - \Delta Var_c^{RURAL} = 8 \text{ pp} - (-16 \text{ pp}) = 24 \text{ pp}$. For the high vs. low school density comparison, the results are 18 *pp* and 17 *pp* respectively. Table 7 summarizes the calculations.

Assumption 3.a states that the change in general classroom assignment practice is the same in areas with different costs of school choice. As argued previously, it is not restrictive, and the actual effect of school competition on sorting within a school should be close to the upper bound estimate (24 *pp* or 17 *pp*). However, Assumption 3.b is unlikely to be true and the mixing effect of larger catchment areas should be bigger in areas with a high cost of school choice. I relax this assumption and claim that the mixing effect is proportional to the ratio of elementary schools to lower-secondary schools. Table 1

¹³ That is, when Assumptions 2, 3.a and 3.b hold.

shows that the ratio for rural areas is 2.31 elementary schools per lower-secondary school, while for urban areas the ratio is 1.49. Table 5 demonstrates that in rural areas sorting between schools drops by 21 *pp* between the two stages of education. Hence, back-of-the-envelope calculations suggest that the mixing effect in urban areas is: $1.49/2.31 = 0.651$ times 21 *pp*, which equals 13.7 *pp*. Based on this, the effect of student mobility on sorting between schools is $15 \text{ pp} + 13.7 \text{ pp} = 28.7 \text{ pp}$ of the proportion of the variance explained by the school level. For areas with a low density of lower secondary schools, the ratio is 3.12, and for areas with a high density of such schools it is 1.81. Consequently, the effect on sorting between is $-2 \text{ pp} + (1.81 / 3.12) * 20 \text{ pp} = 9.6 \text{ pp}$ of the proportion of the variance explained by the school level. The larger effect of school competition in the case of the urban vs. rural comparison is not surprising given that this division is more contrasting than the comparison based on school density (see Table 1).

Table 7. The Effects of Interest

Interpretation	Lower Bound		Upper Bound
	<i>Urban</i>	<i>Rural</i>	<i>Difference</i>
Within	9 pp - 1 pp = 8 pp	6 pp - 1 pp + (5 pp - 26 pp) = -16 pp	24 pp
Between	28 pp - 13 pp = 15 pp	5 pp - 26 w = -21 pp	36 pp
	<i>High LS/Km²</i>	<i>Low LS/Km²</i>	Difference
Within	6 pp - 1 pp + (17 pp - 19 pp) = 3 pp	8 pp - 2 pp + (5 pp - 25 pp) = -14 pp	17 pp
Between	17 pp - 19 pp = -2 pp	5 pp - 25 pp = -20PP	18 pp

Notes: The table presents the logic behind the lower and upper bound estimates of the effect of school competition on sorting between schools and within a school. The numbers used in calculations come from Table 5.

Source: author's own elaboration.

Robustness

Test-room shocks at the time of measurement could artificially lead to more homogeneous schools or classes. To see this, suppose that a barking dog was influencing students' attention during the test, inducing a correlation between the test scores of students from the same test room and thus a measure of their homogeneity. This is especially problematic in the EVA survey, as the students from lower-secondary schools took Raven's test in groups, while those from elementary schools took it separately. This difference would imply more homogeneous classes in the former. However, there are three reasons why this scenario is unlikely. First, the team of professional psychometricians conducted the measurement with all the measures taken to provide a neutral environment for the test takers [Jasińska, Hawrot, Humenny, Majkut, Konlewski, 2013]. Second, the nature of these shocks would have to be different between areas with different costs of school choice, which is unlikely. Third, to fully exclude

this possibility, I exploit the fact that in almost one-third of the lower-secondary schools, students took Raven's test in two groups within a class. Thanks to this, I can directly test whether there is any impact of being in a separate group on Raven's score after controlling for class fixed effects. Any significant effect would indicate that the test-room environment matters for the outcome; the coefficients are, however, highly insignificant across areas with different costs of school choice. On the other hand, the correlation between a student's Raven's score and the average of his/her classmates from the same testing group is significantly higher than the correlation with the other group (from the same class). Nevertheless, the difference is larger in rural areas, which is against this alternative explanation (the results are available upon request).

The difference in sorting across the stages of education might reflect the cohort-specific differences in sorting at the entrance to elementary education. For the seventh graders (from 2010), the sorting at their first grade (in 2004) could be different than for the first graders in 2010. To assess this explanation, I compare the share of the seventh and first graders who attended a local elementary school. Table 8 shows a falling trend in local elementary school attendance. In the total sample, the seventh graders were more likely to go to their assigned schools by almost 3 *pp*. The difference is not statistically different from zero in the subsamples. In areas with a low cost of school choice, the difference is somehow higher, 4.8 *pp* for the urban areas and 4 *pp* for regions with high school density. Even though this effect could bias downward the results, its magnitude and significance cast doubts on the importance of it.

As for sorting within a school, there are no reasons why principals' practice could change between 2004 and 2010. The results presented in Table 5 show that sorting within is negligible at the entrance to elementary education. Moreover, there was no institutional change which would have provided additional motivation for student grouping or mixing. Finally, the potential confounding effect would have had to affect sorting differently across areas with different costs of school choice. I find this possibility rather unlikely.

Table 8. Attendance at Local Elementary School

Stage	All	Urban	Rural	High LS/km ²	Low LS/km ²
Elementary school	79.1%	72.4%	82.1%	76%	81%
Lower-secondary school	82%	77.2%	83.4%	80%	84%
Difference	2.9 <i>pp</i> **	4.8 <i>pp</i>	1.3 <i>pp</i>	4 <i>pp</i>	3 <i>pp</i>
N	10 528	3455	7073	5218	5141

Note: Columns (Urban) and (Rural) show the statistics for rural and urban schools, where the urban schools are in municipalities with population larger than 50,000. Columns (Low LS/km²) and (High LS/km²) are for areas with a density of lower secondary schools per km² below and above its median. Percentage of "yes" responses to for the question asked of parents about whether their child attended a local and assigned elementary school. *** denotes significance at the 1% level, ** at the 5% level.

Source: Author's calculation based on the EVA survey.

Explaining Sorting Within a School

The results presented in the previous section show that higher student mobility leads to higher sorting of students between schools and within a school. While the former effect has been extensively investigated [Epple, Romano, 1998; Ladd, Fiske, 2000; Hsieh, Urquiola, 2006; Nechyba, 2006; Böhlmark et al., 2015], there has been little research on sorting across classes. This section briefly reviews the existing studies and explores the underlying mechanisms.

The empirical literature on the determinants of sorting across classes is limited. Card and Rothstein [2007] show that those US schools that are more racially integrated are also more likely to sort students across classes, suggesting that within-school sorting is used as a substitute for between-school segregation. Kalogrides et al. [2013] document that class composition in US schools is far from random and might be detrimental for lower achievers, as they also receive fewer resources (e.g., novice teachers). Collins and Gan [2013] report large variations in classroom assignment practices, but they also show that grouping students together might be beneficial for low- and high-skill students because of tailoring teaching practice to the skills of students.¹⁴

Two theoretical works directly link student mobility and school competition with the sorting of students. Epple et al. [2002] argue that the creation of a high track within a school might be used to attract high-skill or high-income students (demand for peer quality), while Clotfelter et al. [2005] suggest that this could be used to attract highly skilled teachers (demand for teachers). Suppose that students differ by skill and maximise the expected difference between the benefits and costs of education. The benefits are a function of peer quality and teacher skills, whereas the costs depend on the distance to a school. Students select a non-local school only if peer and teacher quality effects overbalance the extra costs of a longer travel distance. Next, suppose that school principals maximise enrolment and have to accept all local students. Because the number of students depends on the expected benefits from education, principals also indirectly care about school quality. They decide whether to sort or mix students across classes. Sorting yields administrative costs and requires adjusting of teaching practices. Teachers differ in their skills, and they select a school that maximises their utility, which is an increasing function of wage (fixed across schools) and classroom environment (determined by the quality of students).

When the cost of school choice is high enough (e.g. it is expensive to travel), students never select an alternative school, and the school principals have no motivation to introduce within-school tracking. When the school choice is feasible, students are more likely to choose a non-local school if they live in a low-quality area. Consequently, a school from a low-quality area has to pro-

¹⁴ For a critical review of this paper, see Burriss and Allison [2013]. The results are consistent with Duflo, Dupas and Kremer [2011], who find positive effects of randomly assigned within-school tracking.

vide skilled teachers or high-quality class peers to keep local high-achievers and attract non-local ones. In other words, school choice, together with residential sorting, might motivate principals to use classroom sorting as a means of competition for high-skill students (the demand for peer quality channel). Also, since teacher wages are fixed, the only way to attract skilled teachers is by offering them a pleasurable teaching environment (the demand for teachers channel).

Table 9. Lower-Secondary Principals

	Urban	Rural	Difference
<i>Panel A: Principals and the External Examination</i>			
6 th grade exam as a good signal	67.2%	55.6%	11.4 pp
Usage of the 6 th grade exam	84.8%	77.8%	7 pp
External examination as a good signal	93.5%	83.4%	9.9 pp**
External examination is random	18%	26.3%	8.3 pp
External examination is too influential	62%	47%	15 pp
<i>Panel B: Principals' characteristics</i>			
Experience in schooling (years)	24	24.3	0.3
Experience as a principal (years)	11.2	9.9	1.3
% of females	70%	60%	10 pp
N	46	104	
	High LS/km ²	Low LS/km ²	Difference
<i>Panel C: Principals and the External Examination</i>			
6 th grade exam as a good signal	64.3%	52.5%	11.8 pp
Usage of the 6 th grade exam	77.3%	80.8%	-3.5 pp
External examination as a good signal	87.9%	83.4%	4.5 pp
External examination is random	25.5%	23.8%	1.7 pp
External examination is too influential	64.9%	36.6%	28.3 pp***
<i>Panel D: Principals' characteristics</i>			
Experience in schooling (years)	23.9	24.6	0.7
Experience as a principal (years)	10.9	9.5	1.4
% of females	61.8%	61.8%	0 pp
N	77	73	

Note: Urban and Rural show the statistics for rural and urban schools, where urban schools are in municipalities with a population larger than 50,000. Low LS/km² and High LS/km² are for areas with a density of lower-secondary schools per km² below and above its median. The "6th grade exam as a good signal" variable is an answer to the question "Is the 6th grade exam a good measure of skills of students who are attending your school?"; "Ext. exam as a good signal" is an answer to "Do you agree that an external examination makes it possible to compare students' achievements?"; "Ext. exam is random" is an answer to: "Do you agree that the examination scores are pretty much random?"; "Ext. exam is too influential" is an answer to: "Do you agree that the examination scores matter too much in the educational path of a child?" All of the above variables equal one for: "strongly agree"/"rather agree" and 0 for "rather disagree"/"strongly disagree". The "Usage of the 6th grade exam" variable is one if a principal's school analysed examination scores and used them somehow. *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level.

Source: author's own elaboration.

The first part of this section presents the survey data on lower-secondary school principals' characteristics and their sorting practices. The second part empirically evaluates the two theoretical mechanisms: demand for peer quality and demand for teachers. The results suggest that demand for peer quality is the most likely explanation for the positive effect of student mobility on sorting within a school.

Survey Data on School Principals

The EVA survey includes an open question about class assignment.¹⁵ In general, principals from all types of areas underline the equal distribution of high and low achievers across classes. This practice stands in contrast with the findings of this paper. In 2010, there was strong political pressure to equalise educational opportunities, and thus principals might do not want to speak about their sorting practices openly. On the other hand, the political pressure can explain why students are more mixed across classes when entering lower-secondary education in rural or low school density areas.

The attitudes and characteristics of lower-secondary school principals might shed light on the reasons behind an increase in sorting within a school. Table 9 presents the results for 150 lower-secondary schools. Panels A and C show that principals from urban and high-school-density areas are more likely to trust and use external examinations. But they believe that the score matters too much in the educational path of a child. These results are consistent with the observed higher sorting across classes and schools. Differences in the principal's characteristics might matter as well. However, as Panels B and D show, principals across the studied areas have almost identical work experience,¹⁶ and they are equally likely to be female (except that the share of females is higher in urban areas).

Demand for Peer Quality

School principals might decide to create a high track within their schools to attract non-local students or to keep local ones. For each lower-secondary school in the EVA survey, I correlate a measure of student sorting based on Raven's score with a measure of sorting based on the non-locality of students.¹⁷ If the demand for peer quality is a driving force for the homogeneity of classes, there should be a positive association between the two measures of sorting. Since the focus is on sorting at the entrance to lower-secondary education, the observations from elementary education are excluded. In particular, I follow Collins and Gan (2013) and define the sorting of students across classes as:

¹⁵ The reliability of this kind of data is discussed in Betts and Shkolnik [2000].

¹⁶ Because of hiring criteria all principals have the same level of education.

¹⁷ Unfortunately, the available data do not allow us to check whether school principals use sorting to keep local high-achievers.

$$W_s^R = \frac{1}{2} \sum_c \frac{\hat{\sigma}_{cs}^R}{\hat{\sigma}_s^R}$$

where $\hat{\sigma}_{cs}^R$ is the observed standard deviation of Raven's score for class c from lower secondary school s and $\hat{\sigma}_s^R$ is the observed standard deviation of Raven's score for lower secondary school s . The ratio is at class level, but I calculate the school-level average (there are two classes per school). In the case of perfect sorting across classes, the class-level variance is zero, but the school-level variance is positive. Hence the measure W_s^R is null. With perfect mixing, the variance at class level is the same as at school level and W_s^R is one.¹⁸ I define a similar measure for sorting based on the non-locality of students:

$$W_s^N = \frac{1}{2} \sum_c \frac{\hat{\sigma}_{cs}^N}{\hat{\sigma}_s^N}$$

where $\hat{\sigma}_{cs}^N$ is the class-level observed standard deviation of a dummy indicating whether a student is non-local and analogously $\hat{\sigma}_s^N$ is for school level. The regression of interest is:

$$W_s^R = \alpha + \beta W_s^N + \epsilon_s$$

Table 10 Columns (3) and (7) show that switching from perfect sorting to mixing in urban areas increases Raven's sorting measure by .254 on average, and by .052 in regions with dense school networks, implying more heterogeneous classes. Columns (5) and (9) show that in rural and low-density areas, the correlations are negative and insignificant. These results are consistent with the demand for peer quality hypothesis.

The measure of sorting based on the non-locality of students might be misleading when there are only few non-local students. To see this, suppose that there is one non-local student and she is randomly assigned to a class. The measure for the assigned class will be one, while being zero for the second class. Consequently, the school-level average measure of sorting will be half, even though the non-local student was assigned randomly. The absolute difference between classes in the share of non-local students is an alternative measure of sorting based on non-locality. Since the EVA survey contains data on two classes per school, the measure is defined as $|NonLocal_{1s} - NonLocal_{2s}|$, where $NonLocal_{1s}$ is the share of non-local students in the first class from school s . The regression of interest is:

$$W_s^R = \alpha + \beta |NonLocal_{1s} - NonLocal_{2s}| + \epsilon_s$$

¹⁸ The sorting measure can be larger than unity. This might happen when one class consists of students from the middle of the distribution and the second class includes students from the bottom and top of the distribution.

An increase in $|NonLocal_{1_s} - NonLocal_{2_s}|$ implies higher sorting across classes based on the non-locality of students. Consequently, the demand for peer quality hypothesis implies a negative correlation in urban and high-density areas and null in rural and low-density areas. Table 10 Columns (4) and (8) show that introducing complete segregation increases sorting based on Raven's score by .209 on average for urban areas and by .128 for high-density areas (the negative coefficients imply more sorting). The coefficients for the other regions are not statistically significant from zero.

Demand for Teachers

High-quality teachers can improve the attractiveness of a school. However, because teachers' wages cannot vary across public schools in Poland, principals cannot use this margin to attract skilled teachers. Instead, they might decide to create homogeneous and high-track classes, which are easier to teach. To test for this possibility, I correlate teacher characteristics with the class average of Raven's score and control for school fixed effects. A positive association between the measures of teacher experience and the class-level average of Raven's score would be consistent with the demand for teachers hypothesis. The focus is only on teachers and classes from lower-secondary education. The regression of interest is specified as follows:

$$\bar{Y}_{cs} = \alpha + \beta T_{cs} + \mu_s + \epsilon_{cs}$$

where the dependent variable is the average Raven's score for class c from lower-secondary school s , T_{cs} is the class-level average of teacher characteristics, and μ_s is the school fixed effects. I use two measures of teacher characteristics: teaching experience in years and the professional rank. The rank ranges from novice teacher (=0) to "professor of education" (=5).

Table 10 Columns (3) and (7) document no significant correlation between the class-level averages of the teacher's rank and Raven's score for urban and high-density areas. On the one hand, there is no significant effect for rural areas (Column (5)), but the effect is significant and negative for areas with low school density (Column (9)). An increase in the average teacher's rank by one grade reduces the average class's Raven's score by -.287 of standard deviation. Columns (4), (6), (8) and (10) show the same regression, but with teaching experience as an additional independent variable. The magnitude of the coefficient on the teacher's rank doubles for urban areas but remains insignificant. The coefficient on teaching experience is not significant and close to zero.

Table 10. Demand for Peer Quality and Demand for Teachers

	All		Urban		Rural		High LS/km ²		Low LS/km ²	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Panel A: Dependent Variable: Sorting based on Raven's Score (W_s^R)</i>										
Sorting based on Non-Locality (W_s^N)	.010 (.054)		.254 (.070)***		-.055 (.055)		.052 (.071)		-.030 (.076)	
$ NonLocal_{i,s} - NonLocal_{2s} $		-.096 (.046)**		-.209 (.052)***		.021 (.047)		-.128 (.048)***		-.030 (.109)
Constant	.948 (.050)***	.970 (.008)***	.699 (.068)***	.987 (.014)***	1.013 (.050)***	.959 (.009)***	.912 (.066)***	.983 (.009)***	.983 (.070)***	.957 (.015)***
N	136	141	41	44	95	97	69	72	67	69
R ²	.0005	.043	.214	.268	.019	.002	.012	.119	.005	.002
<i>Panel B Dependent Variable: Mean of Raven's score at the class level</i>										
Teacher's Rank	-.080 (.080)	-.096 (.101)	-.025 (.173)	-.223 (.222)	-.100 (.087)	-.081 (.111)	.099 (.093)	.028 (.146)	-.287 (.101)***	-.239 (.126)*
Teacher's Experience		.002 (.008)		.026 (.019)		-.003 (.009)		.009 (.012)		-.007 (.010)
N	267	266	90	90	177	176	142	142	125	124
R ²	.681	.682	.729	.737	.632	.635	.754	.756	.636	.641

Notes: Panel A presents the results of the OLS regression of the standardised Raven's score sorting measure on the non-locality sorting measure and the distance in the share of non-local students between classes. The measures are calculated for the 7th graders only. The unit of observation is school (lower-secondary). Panel B presents the OLS regressions of the class average of the standardised Raven's score on the class averages of teacher experience in years and the teacher's professional rank. The rank ranges from 0 to 5, where 0 is the lowest ranked teacher and 5 is "the professor of education". The measures are calculated for the 7th graders only. The unit of observation is class (from gimnazjum). The robust standard errors are presented in the parentheses. All the estimations are weighted using the survey weights. *** denotes significance at the 1% level, ** at the 5% level, and * at the 10% level.

Source: author's own elaboration.

The results suggest that lower-secondary school principals do not offer high tracks to attract skilled teachers. If anything, principals might assign higher-rank teachers to worse classes in order to compensate low-peer quality or improve discipline.¹⁹ Regardless of the reasons, this might have a positive effect on the educational equality of opportunity. However, more data is needed to fully investigate this possibility.

Conclusions

This study shows that facilitating student mobility leads to increased sorting of students within a school and between schools. It links sorting across classes with students' demand for peer quality, which motivates school principals to create selective tracks. The results bear relevance for policy makers who wish to use student mobility and school competition as a means to improve the quality of schools, but also want to avoid its negative distributional consequences. The results underline the importance of school principals' incentive structure. Principals might create classes with a high level of peer quality to attract high achievers or high-income students. Within-school tracking could be weakened by the incorporation of value-added estimates of school performance into principals' objectives, as it motivates them to compete for low-background and low-performing students [MacLeod, Urquiola, 2009]. Even though value-added-based accountability has been heavily discussed, not much attention has been paid to potential distributional effects [Rothstein, 2009; Angrist, Pathak, Walters, 2011; Chetty, Friedman, Rockoff, 2014]. An alternative policy could be to use mean-tested school vouchers.²⁰

The results also shed light on the potential distributional consequences of the Polish education reform of 2017–2020. The reform abolished the lower-secondary stage and expanded elementary education from six to eight years. After the 8th grade students now move to a four-year higher-secondary stage. This study shows that the transition from one stage of comprehensive education to the next leads to less (more) intensive sorting of students in rural (urban) areas. This implies that the move from a two-stage to one-stage design of comprehensive education might lead to a rise (fall) of education inequalities in rural (urban) areas. One must be careful not to draw too strong conclusions though, as sorting to elementary school might change as a result of the reform and thus the homogeneity of students in elementary schools in the two-stage system might not be a valid counterfactual for the scenario with the one-stage system.

¹⁹ An alternative explanation is that high-skill teachers prefer to teach low-skill students. However, in this scenario, high-skilled students will not be attracted by teacher quality, and thus school principals have no motivation to hire skilled teachers.

²⁰ This policy is in effect in Chile and the Netherlands, see Böhlmark et al. [2015].

References

- Ajayi Kehinde F. [2012], *School Choice and Educational Mobility, Unpublished Working Paper*, Boston University.
- Angrist J.D., Pathak P.A., Walters Ch.R. [2011], *Explaining charter school effectiveness, Technical Report*, National Bureau of Economic Research.
- Betts J.R., Shkolnik J.L. [2000], Key difficulties in identifying the effects of ability grouping on student achievement, *Economics of Education Review*, 19(1): 21–26.
- Böhlmark A., Holmlund H., Lindahl M. et al. [2015], *School choice and segregation: Evidence from Sweden, Technical Report*.
- Brouwers S.A., Van de Vijver F. JR, Van Hemert D.A. [2009], Variation in Raven's Progressive Matrices scores across time and place, *Learning and Individual Differences*, 19(3): 330–338.
- Burris C.C., Allison K.E. [2013], *Review of "Does sorting students improve test scores?"*, Boulder, CO: National Education Policy Center.
- Card D., Rothstein J. [2007], Racial segregation and the black-white test score gap, *Journal of Public Economics*, 91(11): 2158–2184.
- Chetty R., Friedman J.N., Rockoff J.E. [2014], Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates, *The American Economic Review*, 104(9): 2593–2632.
- Clotfelter Ch.T, Ladd H.F., Vigdor J. [2005], Who teaches whom? Race and the distribution of novice teachers, *Economics of Education Review*, 24(4): 377–392.
- Collins C.A., Li Gan [2013], *Does sorting students improve scores? An analysis of class composition, Technical Report*, National Bureau of Economic Research.
- Dolata R. [2008], *Szkoła, segregacje, nierówności*, Wydawnictwo Uniwersytetu Warszawskiego.
- Duflo E., Dupas P., Kremer M. [2011], Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya, *The American Economic Review*, 101(5): 1739–1774.
- Epple D., Romano R.E. [1998], Competition between private and public schools, vouchers, and peer-group effects, *American Economic Review*: 33–62.
- Figlio D.N., Page M.E. [2002], School choice and the distributional effects of ability tracking: does separation increase inequality?, *Journal of Urban Economics*, 51(3): 497–514.
- Friedman M. [1995], *The role of government in education*, Rutgers University Press.
- Friedman W.H., Lucas A.M. [2017], The importance of information targeting for school choice, *American Economic Review*, 107(5): 638–43.
- Gibbons S., Machin S., Silva O. [2008], Choice, competition, and pupil achievement, *Journal of the European Economic Association*, 6(4): 912–947.
- Herbst M., Herczyński J., Levitas A. [2009], *Finansowanie oświaty w Polsce: diagnoza, dylematy, możliwości*, Wydawnictwo Naukowe Scholar.
- Hoxby C.M. [2000], Does competition among public schools benefit students and taxpayers?, *American Economic Review*: 1209–1238.
- Hsieh Ch.-T., Urquiola M. [2006], The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program, *Journal of Public Economics*, 90(8): 1477–1503.
- Instytut Badań Edukacyjnych [2011], *Raport o stanie edukacji 2010*.

- Jasińska A., Hawrot J., Humenny G., Majkut P., Konlewski M. [2013], *Kontekstowy model efektywności nauczania po I etapie edukacyjnym – wyniki badania podłużnego w szkołach podstawowych*, *Technical Report*, Instytut Badań Edukacyjnych.
- Jensen A.R. [1998], *The G Factor: The Science of Mental Ability*, Praeger.
- Jensen R. [2010], The (perceived) returns to education and the demand for schooling, *The Quarterly Journal of Economics*, 125(2): 515–548.
- Kalogridis D., Loeb S., Béteille T. [2013], Systematic sorting: teacher characteristics and class assignments, *Sociology of Education*, 86(2): 103–123.
- Kane T., Stephanie J., Riegg K., Staiger D.O. [2006], School quality, neighborhoods, and housing prices, *American Law and Economics Review*, 8(2): 183–212.
- Kern N., Thukral R., Ziebarth T. [2012], A Mission to Serve: How Public Charter Schools Are Designed to Meet the Diverse Demands of Our Communities. Issue Brief, *National Alliance for Public Charter Schools*.
- Kremer M., Duflo E., Dupas P. [2011], Peer Effects, Teacher Incentives, and the Impact of Tracking,” *American Economic Review*, 101.
- Ladd H., Fiske E. [2000], *When schools compete: A cautionary tale*, Washington, DC: The Brookings Institution.
- Lucas M.E. [2004], What’s in a grade? School report cards and the housing market, *American Economic Review*: 591–604.
- MacLeod W.B., Urquiola M. [2009], *Anti-lemons: school reputation and educational quality*, *Technical Report*, National Bureau of Economic Research.
- Meghir C., Palme M. [2005], Educational reform, ability, and family background, *The American Economic Review*, 95(1): 414–424.
- Nechyba T.J. [2006], Income and peer quality sorting in public and private schools, *Handbook of the Economics of Education*, 2: 1327–1368.
- Newlon E., Romano R. [2002], Ability tracking, school competition, and the distribution of educational benefits, *Journal of Public Economics*, 83(1): 1–48.
- Raven J. [2000], The Raven’s progressive matrices: change and stability over culture and time, *Cognitive Psychology*, 41(1): 1–48.
- Rothstein J. [2009], Student sorting and bias in value-added estimation: Selection on observables and unobservables, *Education*, 4(4): 537–571.
- Szmigel M.K. [2013], *Gimnazjum niejedno ma imię? – na podstawie wywiadów z dyrektorami. Doniesienie z badań jakościowych*, *Technical Report*, XIX Konferencja Diagnostyki Edukacyjnej, Gniezno.
- Tiebout Ch.M. [1956], A pure theory of local expenditures, *The Journal of Political Economy*: 416–424.