# School Competition and Sorting of Students*

Paweł Bukowski †

July 31, 2017

This study estimates the effect of school competition on sorting between and within schools. The identification strategy is based on a two-stage design of the Polish Comprehensive Education, which allows isolating an exogenous change in student mobility. The measurement of sorting of students is based on Raven's Progressive Matrix test score. The results show that school competition leads to a higher sorting of students within a school and between schools. Next, the two mechanisms behind the effect on sorting within a school are tested. Firstly, student's demand for peer quality could motivate school principals to create selective tracks. Secondly, to attract skilled teachers, school principals might offer them, homogeneous classes. The data point to the importance of the former mechanism.

*JEL Classification: I24*

*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 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 the CERGE-EI/GDN Grant RRC16+12. All errors are mine.

†Centre for Economic Performance, The London School of Economics and Political Science, Houghton Street, WC2A 2AE London, the United Kingdom. E-mail: p.bukowski@lse.ac.uk

# 1 Introduction

It has been argued that school competition might motivate school principals to improve school quality (Friedman, 1955; Hoxby, 2000). When school funding depends on the number of students and people care about the quality of education, allowing free school choice will drive away students from their current and low productive school. This process would continue until higher quality schools dominate the whole educational market or 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 and Silva, 2008; Kern, Thukral and Ziebarth, 2012).

There is, however, an efficiency-equality trade-off associated with the expansion of school competition. Advantageous and high-skill students might be more likely to exert their school choice, if the access to information, student mobility or performance depend on parental resources. School principals have incentives to attract only the best or wealthiest so that their schools perform better on rankings. Consequently, rich or high-performing (poor or low-skilled students) will be concentrated in high-quality (low-quality) schools, which might further reinforce their advantage (disadvantage) (Epple and Romano, 1998; Ladd and Fiske, 2000; Hsieh and Urquiola, 2006; Nechyba, 2006; Böhlmark, Holmlund, Lindahl et al., 2015).

Sorting across schools is not the only potential side effect of school competition. There are two theoretical arguments that competition might also lead to sorting within a school. Firstly, the creation of a high track might be used by school principals for "cream skimming" of students (Epple, Newlon and Romano, 2002). Secondly, the creation of homogeneous and easy to teach classes might be used to attract high-skilled teachers (Clotfelter, Ladd and Vigdor, 2005). While numerous studies show that school competition leads to sorting of students between schools, we know very little about the effect on sorting within a school (Card and Rothstein, 2007; Kalogrides, Loeb and Béteille, 2013; Collins and Gan, 2013). 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 faced by students (Meghir and Palme, 2005; Kremer, Duflo and Dupas, 2011; Figlio and Page, 2002).

This paper investigates the effect of school competition on sorting within and between schools. The identification strategy is based on a two-stage design of the Polish comprehensive education. Admission to both stages is based on catchment areas with a school choice option. Students are more likely to exert the choice option at the entrance to the secondary stage, implying higher competition among these schools. However, this is true only in areas with low cost of school choice (e.g. urban areas). Capturing the effect of school competition on inequalities requires thus two steps. First is to compare sorting of students at the entrance across the stages of education, for areas with low cost of school choice. Second, to juxtapose this difference with the counterfactual difference for areas with high cost of school choice. The measurement of sorting of students is based on Raven's Progressive Matrix test score, a measure of general intelligence, which is determined by student genetic abilities and socioeconomic background. It is fixed since early childhood, which ensures that the only source of class/school homogeneity is sorting of students. The results show that school competition increases sorting of students both across and within schools.

Next, the data on school characteristics is used to explore the two mechanisms linking school competition with sorting across classes. The results show that high track might be used to attract high-skill or high-income students, which is consistent with a theoretical model developed by Epple et al. (2002). there is no evidence that school principals attract high 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 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 school competition on sorting across classes. Finally, the seventh section concludes.

## 2 Institutional Background

The Polish comprehensive education is compulsory and consists of six years of elementary school (ISCED 1), which is followed by three years of the lower secondary school called *gimnazjum* (ISCED 2). Elementary school and *gimnazjum* usually serve the same community of students, but they are separated entities, with different managerial and teaching bodies. After finishing the comprehensive part, a student may finish their education or enrol in academic, mixed or vocational higher secondary schools (ICED 3).

The admission process to the elementary school and *gimnazjum* is the same. It is based on catchment areas, which means that every student has a right to attend an assigned local public school. Because there are more elementary schools than *gimnazja*,[1] the catchment area for the latter is usually larger and contains the catchment areas from several elementary schools. Table 1 shows the ratio of elementary schools to *gimnazja* in a rural-urban breakdown and for areas with high and low density of *gimnazja*. In the rural (low density) areas there are on average 2.3 (3.1)) elementary schools per *gimnazjum* and almost 1.5 (1.8) in the urban (high density) areas. As an alternative to the local school, parents may request a place in an under-subscribed non-local school, but without guaranteed seat. The are no universal recruitment rules for non-local students. Each school's policy is determined by a school principal and a recruitment committee, which usually consists of selected teachers and a school psychologist.

The school principals and the recruitment committee determine classroom assignment, and there are no differences across the stages of education. As for *gimnazja*, the most common practice is to create similar classes in terms of student performance, with similar foreign language proficiency or with students from the same neighbourhood (Szmigel, 2013). The elementary schools cannot sort students based on their performance, it is unknown, but they take into consideration gender composition, place of living or the date of birth. Parents have a right to suggest an alternative class assignment. Importantly, the assignment is fixed across grades, subjects and reallocations are allowed only in exceptional cases. The peer composition of classes is thus relatively constant at each stage of education. In 2010 there were no limits on classroom size.[2]

Students are examined by two standardised, externally graded and obligatory examinations. After the elementary school (the 6th grade) they take a low stake exam, which serves mostly an informational purpose. After *gimnazjum* (the 9th grade) students are tested with a high stake exam, which is used for the admission into the higher secondary stage of education. These two tests are the basis for the official educational *value added* measures of *gimnazjum* performance. School funding is not linked to the school performance, but the Ministry of Education publishes rankings on the website. Also, various unofficial rankings publish the average *levels* of elementary school or *gimnazjum* performance.

[1]Most of the elementary schools were constructed during the past 50 years, while *gimnazja* only after 1999. The network of elementary schools thus reflects the past demographic situation and is relatively dense. The network of *gimnazja*, in turn, is more "rational" in the sense that it is better adjusted to the current demographic needs. Also, elementary schools serve younger children for whom distance to a school matter more than for older children.

[2]Since 2015 the rules in elementary schools have been unified and are based on the date of birth with an option for parents to request an alternative assignment. Since 2013 a class in grades I-III can have maximum 25 students.

There are clear economies of scale for school principals. The central government finances all Polish public schools through a subsidy. In theory, this amount should be sufficient to cover all expenditures on education, excluding investments and pre-school education. In practice, however, it covers only around 50-70% of the costs (Herbst, Herczyński and Levitas, 2009; Instytut Badań Edukacyjnych, 2011) and the rest is covered by local governments. Since the governmental subsidy is stuck to the student (the money goes with her), school funds depend on enrollment. In addition, the school principals of larger schools have more bargaining power when securing additional funds from the local governments. In general, public schools do not advertise themselves, but they can use other ways of signalling their quality. Because the rankings based on levels are more popular than the official estimates of the educational value added, schools might be tempted to improve their position by "cream-skimming" of students. This paper, argues that sorting across classes might be one way of attracting high-performing or rich students.[3]

The local governments determine the teacher salaries and employment conditions in compliance with the universal collective bargaining agreement (*Karta Nauczyciela*). It specifies the minimum level of wage for each teacher's rank.[4]. Also, teachers are eligible for overtime pay, monetary awards and other non-monetary benefits, for instance, accommodation in school's social apartments.

[3]On the other hand, mixing students across classes might be preferred by "egalitarian" school principals or policymakers as it improves educational equality of opportunity.

[4]In 2015 the minimum monthly gross wages ranged from 1513 PLN (340 EUR) to 3109 (700 EUR). Additionally, the average *total* gross salary for each teacher's rank within the municipality must be at least as large as specified in *Karta Nauczciela*. In 2015 these averages ranged from 2717 PLN (612 EUR) to 5000 PLN (1126 EUR).

Table 1: Descriptive Statistics for Municipalities

| Variable | Rural (1) | Urban (2) | Low $Gim/Km^2$ (3) | High $Gim/Km^2$ (4) |
|---|---|---|---|---|
| *Total Numbers* | | | | |
| Elementary Schools | 10894 | 2248 | 4759 | 8321 |
| *Gimnazja* | 5371 | 2071 | 1634 | 5808 |
| *Averages for Municipalities* | | | | |
| Elementary School per *Gimnazja* | 2.31 | 1.49 | 3.12 | 1.81 |
| Elementary School per $km^2$ | 0.05 | 0.25 | 0.03 | 0.09 |
| *Gimnazjum* per $km^2$ | 0.03 | 0.24 | 0.01 | 0.08 |
| Children per Elementary School | 153 | 337 | 129 | 191 |
| Children per *Gimnazjum* | 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 | 10067 | 156004 | 6983 | 24096 |
| Number of Municipalities | 2386 | 93 | 1234 | 1234 |

*Source*: the Central Statistical Office of Poland and Herczyński & Sobotka (2013). *Note*: Columns (1) and (2) show the descriptive statistics for the rural and urban municipalities in Poland, where the urban municipalities are with population larger than 50 000. Columns (3) and (4) are for the areas with the density of *gimnazja* per $km^2$ below and above its median. All numbers are for 2010, except Tertiary Education Share (2002) and Public Transportation $km^2$ (2007).

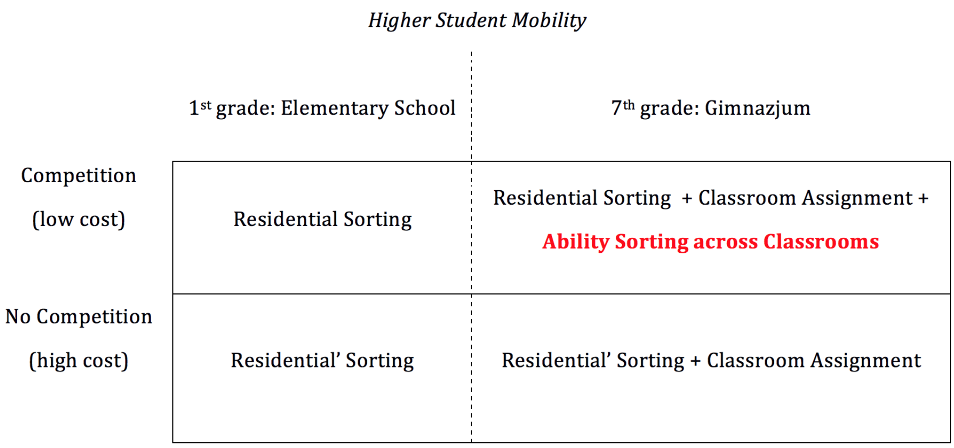


Figure 1: The Identification Strategy - Sorting Within a School

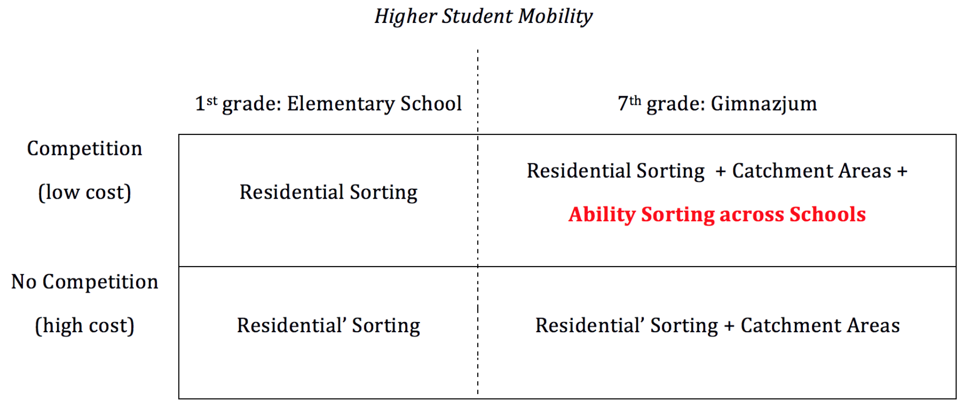


Figure 2: The Identification Strategy - Sorting Between Schools

# 3 Identification Strategy

The effect of school competition on student sorting might be confounded with other parallel social processes. Similar people tend to live together, for instance, because of neighbourhood characteristic, local economic conditions or housing prices (Tiebout, 1956). Local school characteristics influence the latter, which further reinforces self-selection (Figlio and Lucas, 2004; Kane, Riegg and Staiger, 2006). Consequently, the effect of competition on sorting will be biased if schools use catchment areas and school competition coincides with residential sorting.

The identification strategy is based on the two-stage design of the Polish comprehensive education. A comparison of student sorting across classes/schools at the entrance to *gimnazjum* with sorting at the entry to elementary school cancels out the influence of residential sorting and other stage-invariant local characteristics. The difference is then an outcome of the change in student mobility and stage-variant changes in classroom/school assignment. However, in the areas with the high costs of school choice, the difference in student mobility will be irrelevant. Therefore, to isolate the effect of student mobility (school competition) it is sufficient to compare how sorting differs across stages of education, across places with different costs of school choice. The identification strategy is summarized in Figures 1 and 2. Each cell lists the forces driving the classroom (Figures 1) and school (Figure 2) homogeneity across stages of education and locations with the high and low cost of student mobility (potential for school competition). The design is an example of the difference in differences technique. “Treatment” is a change in school competition induced by the change in student mobility. “Treatment group” is an area with the low school choice (high potential for a school competition, the first rows). "Before and after" are the first and second stages of the Polish comprehensive education respectively (the columns). This strategy produces a causal effect of school competition on student sorting under the three assumptions:

**Assumption 1.** ***Treatment*** - *students who enter gimnazjum (the seventh grade) are more likely to use school choice than students who enter the elementary school (the first grade).*

Three observations motivate this assumption. Firstly, students, who enter *gimnazjum* are older and travelling to an alternative school is more feasible for them. Secondly, their performance is known, unlike those who enter the elementary school. Lesser informational constraints might motivate students to select a non-local school and allow school principals to select applicants based on their performance. Thirdly, a catchment area of one *gimnazjum* usually contains catchment areas of several local elementary schools. Consequently, students entering *gimnazjum* are facing larger catchment areas and the school composition of their local *gimnazjum* will to a lesser extent reflect the residential composition of their neighbourhood.

The higher mobility of secondary school students is documented in Table 2. It reports the share of students from the first grade of *gimnazjum* learning in a non-local school and the share of them, which attended a non-local elementary school. The data comes from the Educational Value Added Team survey (described in Section 3). In the whole sample, 18% of students went to a non-local elementary school, and 24% went to a non-local *gimnazjum*. The difference is highly significant. Moreover, in the next section, Table 4 Column (4), I show a suggestive evidence that parents of students entering the second stage might be facing the lesser informational constraints.[5]

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

Table 2: Share of students in non-local schools

| Stage | All | Urban | Rural | Low $Gim/Km^2$ | High $Gim/Km^2$ |
|---|---|---|---|---|---|
| Elem. School | 18% | 23% | 15% | 19% | 16% |
| *Gimnazjum* | 24% | 42% | 16% | 31% | 17% |
| Difference | 6pp*** | 19pp*** | 1pp | 12pp** | 1pp |
| N | 4907 | 1524 | 3383 | 2540 | 2367 |

*Source*: the author's calculation based on the EVA survey. *Note*: Columns (Urban) and (Rural) show the statistics for the rural and urban schools, where the urban school are in municipalities with population larger than 50 000. Columns (Low $Gim/Km^2$ ) and (High $Gim/Km^2$ ) are for the areas with the density of *gimnazja* per $km^2$ below and above its median. All numbers are for 2010. *** denotes significance at the 1% level, ** at the 5% level.

**Assumption 2.** ***Control group*** *- the difference in student mobility across educational stages is irrelevant in areas with the high costs of school choice (low potential for school competition).*

The second assumption is necessary to construct a control group - a group of municipalities without a change in student mobility. In particular areas school variety is limited and the cost of attending a non-local school high (Dolata, 2008). The cost of school choice includes, for instance, a transportation cost, missing links with local peers or more difficult coordination with other parents. Consequently, even though students are more likely to exert their school choice, it is either not attractive or too costly. I use two definitions of groups with the high cost of school choice: 1) rural areas and 2) areas with the number of *gimnazja* per $km^2$ below its median. Table 1 provides descriptive statistics for these groups.The rural municipalities,[6] compared to the urban municipalities, have three (ten) times sparser network of elementary schools (*gimnazja*), twenty-three times sparser network of public transportation and ten times smaller population density. The areas with below the median density of secondary schools, compared to the areas above the median, have also on average higher costs of school choice, but the difference is much smaller. As such this measure provides a useful robustness check since it captures areas, which are relatively similar, except the variety and access to *gimnazja*.

**Assumption 3.a.** ***Common trend*** *(for sorting within a school) - in the absence of an increase in school competition, a change in class assignment between elementary schools and gimnazja is the same in areas with different cost of school choice.*

[6] Those with a population lower than 50 000

**Assumption 3.b.** ***Common trend*** *(for sorting between schools) - the change in the size of catchment areas between elementary schools and gimnazja leads to the same level of between-school student mixing in areas with different cost of school choice.*

The third identifying assumption is an analogue of the "common trend" assumption. It is defined separately for the analysis of sorting within a school and between schools. Assumption 3.a says that reasons to sort or mix students across classes, which are unrelated to school competition, should be similar in areas with different cost of school choice. The qualitative evidence discussed in Section 6.2 supports this presumption. Assumption 3.b is analogous 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 the rural areas as there are more elementary schools per *gimnazjum* than in the urban areas (see Table 1). In other words, the inter-stages difference in school catchment areas will mechanically lead to student mixing or sorting. Section 5.2 replaces this assumption with another the mixing effect is proportional to the ratio of elementary schools to *gimnazja*.

# 4 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 a similar measure for the within-school sorting. The third part presents the data.

## 4.1 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 $\mu$, the school-level deviation from that mean $u_s$, the class-level deviation from the school mean $u_c$ and the residual component $e_{ics}$:

$$y_{ics} = \mu + u_s + u_c + e_{ics} \tag{1}$$

By construction, the variance of the SEB variable at stage $t$ (either Gimnazjum - *gim* or Elementary School -*es*) is a sum of the variance of the school-level component, the variance of the class-level component and the residual variance:

$$Var_t = Var_{s,t} + Var_{c,t} + Var_{e,t} \tag{2}$$

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

$$\Delta Var_s = \frac{Var_{s,gim}}{Var_{gim}} - \frac{Var_{s,es}}{Var_{es}} \tag{3}$$

## 4.2 Sorting Within a School

The change in sorting within a school can be captured similarly, except that one has to correct for the differences in catchment areas between the elementary school and *gimnazjum*. An intensity of sorting within a school is defined as a ratio of the class-level variance to the total variance $\frac{Var_{c,t}}{Var_t}$ . Ignoring the catchment area problem, the change between educational stages is simply $\frac{Var_{c,gim}}{Var_{gim}} - \frac{Var_{c,es}}{Var_{es}}$.

The problem arises because the catchment areas are larger for *gimnazja* than for the 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 and the fraction explained by the class-level increases correspondingly. To see this, suppose that there is just one class per elementary school and students have the same classmates from both elementary school and *gimnazjum*. Because of the nested catchment areas, students from several elementary schools will go to one *gimnazjum*, and each class in that *gimnazjum* will consist of students coming from the same elementary school. The relative importance of the class-level ($\frac{Var_{c,t}}{Var_t}$) increases, even though there is no change in student sorting across classrooms.[7] To correct for this problem one can adjust for the negative change in the fraction of the variance explained by the school-level. I propose the following measure of the change in sorting within a school:

$$\Delta Var_c = \frac{Var_{c,gim}}{Var_{gim}} - \frac{Var_{c,es}}{Var_{es}} + \mathbb{1}_{[\Delta Var_s < 0]} \Delta Var_s \tag{4}$$

where $\mathbb{1}_{[a]}$ 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 the school-level is negative. Intuitively, the aforementioned problem arises only when *gimnazja* have

[7]The way of looking at this problem is to realise that, in this scenario, schools at the elementary school stage become classes at the *gimnazjum* stage. With one class per elementary school, there is no difference between labels: "school" and "class". Although there is no change in the class composition at the transition to *gimnazjum*, the distinction between "school" and "class" begins to matter. This is because of groups of students, which were "classes/schools" at the elementary stage, becomes "classes" at the secondary stage.

larger catchment areas than the elementary schools and their ratio $\frac{Var_{s,t}}{Var_t}$ is lower. When there is no change in the class composition, but catchment areas are larger for the secondary schools, $\frac{Var_{c,gim}}{Var_{gim}} - \frac{Var_{c,es}}{Var_{es}} = -\Delta Var_s$ and thus $\Delta Var_s$ should be 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 the class-level captures the effect of interest.

To estimate the proportions of variance explained by the class and school levels, I use a multilevel mixed-effects linear regressions (also called a hierarchical linear model).

Comparing changes in sorting across areas with the different cost of school choice (potential for school competition) isolates the effect of school competition. I consider two binary[8] measures: urban versus rural, and the number of *gimnazja* per $km^2$. above the median versus below. The underlying assumption is that in the rural areas and the areas with school density below the median, the cost of school choice is so high that everybody follows their local school.

The effect of school competition on sorting within and between schools is defined as:

$$\Delta Var_c^{Comp} - \Delta Var_c^{NoComp} \tag{5}$$

$$\Delta Var_s^{Comp} - \Delta Var_s^{NoComp} \tag{6}$$

## 4.3 Data

The data are drawn from the sample of Polish students collected by the Educational Value Added Team.[9] The cross-section is from 2010 and consists of 5600 first-graders and 5567 seventh-graders (which is an entry grade of *gimnazjum*) from 330 randomly drawn public schools in Poland.[10] A set of student, parental, teacher, school and municipality - level characteristics is available. Importantly, it includes questions about each school's sorting practices. All the statistics used in the paper are weighted using the survey weighting scheme. The results should be interpreted as representative of the Polish population. Table 3 summarizes the available sample.

[8]Continuous variables are not feasible. There are no established methods for estimating the effect of a continuous variable on a change in the composition of a variance of another variable.

[9]The Project was funded by the European Union under the European Social Fund and was ran by the Central Examination Board until September 2012. Since October 2012 the project is run by Educational Research Institute in Warsaw.

[10]The target population were elementary public schools with first grades larger than ten students and public *gimnazja* with seventh grades larger than 20 students.

The main outcome variable and a measure of student's socioeconomic characteristics is a standardized (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, p.2). In other words, eductive and reproductive abilities allow to understand concepts and learn new material. They are components of an underlying general mental ability (Jensen, 1998). The test usually consists of 4x4 3x3 or 2x2 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 why students might have a similar level of Raven's score is self-selection. The advantage of Raven's score is that it includes characteristics affecting sorting of students, such as genotype, which are not necessarily captured by other measures (e.g., mother's education).

I test the claim that Raven's score is not affected by education by regressing mother's and father's education on Raven's score, a dummy denoting observations from *gimnazjum* (the seventh grade) and an interaction term between the two. If education (and other inputs) between elementary school and *gimnazjum* do not matter for Raven's score, there should be no difference in the correlation between parental education and Raven's score for the first and seventh graders. Table 4 Columns (1) and (2) show that while there is a positive correlation between mother's/father's education and Raven's score, it is not significantly different across the grades.

[11] For each student $i$ from grade $g$, I calculate Raven's z-score $zscore_{ig} = \frac{score_{ig} - \overline{score_g}}{sd(score_g)}$ where $score_{ig}$ is raw Raven's score and $sd(score_{ig})$ is a standard deviation of Raven's score for each grade.

## Table 3: Descriptive Statistics of the Sample

| | Elementary School | | | | | *Gimnazjum* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Obs. | Mean | St. Dev. | Min | Max | Obs. | Mean | St. Dev. | Min | Max |
| *Full sample* | | | | | | | | | | |
| 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 | | | | |
| *Urban* | | | | | | | | | | |
| 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 | | | | |
| *Rural* | | | | | | | | | | |
| 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 | | | | |
| *Low* $Gim/Km^2$ | | | | | | | | | | |
| 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 | 6 | 28 |
| Numbef of schools | 99 | | | | | 73 | | | | |
| *High* $Gim/Km^2$ | | | | | | | | | | |
| 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 |
| Numbef of schools | 81 | | | | | 77 | | | | |

*Note*: Urban (rural) schools are in municipalities with population larger (smaller) than 50 000. High (low) $Gim/Km^2$ schools are in municipalities with the density of *gimnazja* per $km^2$ above (below) and above its median. All numbers are for 2010. Unweighted statistics

Table 4: Raven's Score and Education

| *Dependent Variable:* | Mother's Education (1) | Father's Education (2) | Desired Education for a Child (3) | 6th grade GPA (4) |
|---|---|---|---|---|
| Raven's Score | .557 | .543 | .464 | .532 |
| | (.042)*** | (.040)*** | (.035)*** | (.017)*** |
| *Gimnazjum* | -.265 | -.219 | -.352 | |
| | (.072)*** | (.072)*** | (.065)*** | |
| Raven's Score *X Gimnazjum* | -.019 | -.008 | .370 | |
| | (.051) | (.050) | (.042)*** | |
| N | 10320 | 10167 | 10376 | 4896 |
| Estimator | OLogit | OLogit | OLogit | OLS |

*Notes*: The table shows regressions of the depended variables on the standardized Raven's Progressive Matrix Test score, a dummy indicating observation from the seventh grade - *Gimnazjum* (excluded category is the first grade - elementary school), and the interaction between them. Mother's and Father's Education are categorical variables, which take valuess between 1 and 9, where 1 is unfinished elementary education and 9 is PhD. Desired Education for a Child is a categorical variable, which takes values between 1 and 7, where 1 is vocational education and 7 PhD. 6th grade GPA is the average of grades from various subjects, it ranges between 2 and 6, where 2 is the worst. 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.

On the other hand, Column (3) shows that there is a more positive correlation between Raven's score and desired education for a child in the seventh grade than in the first grade. The positive coefficient is consistent with the lesser informational constraints faced by parents at the entrance to *gimnazjum*. Since, as reported in Column (4), there is a positive correlation between the sixth grade GPA and Raven's score, students with 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 to the elementary school, and the correlation between Raven's score and the desired education is significantly lower.

# 5 Results

The first part of this section presents the decomposition of the variance of Raven's score and translates it into the effect of school competition on the sorting of students. The second part shows the robustness checks.

## 5.1 Decomposition of the Variance of Raven's Score

Table 5 presents the proportion of variance of Raven's score explained by the school and the class levels, in a breakdown by the stages of education, and by the urban and rural areas. Table 6 presents similar estimates for the areas above and below of the median of *gimnazja* 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.

In the urban areas, the school level explains 13%, and the class level explains 1% of Raven's score variation at the entrance to the elementary school. At the entrance to *gimnazjum*, the proportions increase to 28% and 9% respectively. It means that *gimnazja* and *gimnazja*'s classes are more homogeneous than in the case of the elementary school. Consequently, the explained proportion of variance grows from 14% to 37%. The same pattern, but smaller in magnitude, is documented for the 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 student's increased student mobility (higher school competition) and other grade-variant changes in the school and class assignment (Assumption 1). To isolate the former mechanism one needs to compare this difference with the difference for the control areas without the change in student mobility, that is, for the rural areas or the areas with secondary school density below its median (Assumptions 2, 3.a and 3.b). At the entrance to the elementary school, the school and class levels explain 26% and 1% of Raven's score variation in the rural areas. In the areas with low school density, the school and class levels explain 25% and 2% respectively. At the entrance to *gimnazjum*, the school level drops to 5% in both areas, which means that *gimnazja* are more heterogeneous than elementary schools. Th drop is likely to be explained by the differences in catchment areas sizes. At the same time, the fraction explained by the class level rises to 6% in the rural schools and 8% in the areas with low school density. Interpretation of this change, however, is less straightforward. Suppose that there is just one class per elementary school and students have the same classmates in elementary school and *gimnazjum*. Because of the nested catchment areas, students from several elementary schools will go to one *gimnazjum* and each class in the *gimnazjum* will consist of students coming from the same elementary school. The importance of the class level increases, even though there was no change in the 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 variance, which means that classes are more heterogeneous at the entrance to *gimnazjum* than to the elementary school. Based on Equation 4, the drop in sorting within a school is 16pp for the rural and 14pp for the low school density areas.

Table 5: Decomposition of the Variance of Raven's Score - Urban vs. Rural

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

*Notes*: The table shows decomposition of 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 survey weighting scheme. Urban (rural) schools are in municipalities with population larger (smaller) than 50 000.

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

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

*Notes*: The table shows decomposition of 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 survey weighting scheme. High (low) $Gim/Km^2$ schools are in municipalities with the density of *gimnazja* per $km^2$ above (below) and above its median.

## 5.2 School Competition and Sorting of Students

This section interprets the above results. Ignoring the control group,[12] the difference between the stages of education for the areas with low cost of school choice (high competition) is interpreted as a lower bound of the potential effect of competition on sorting. The reason is that it ignores the mixing effect of catchment areas and general attempts to equalise classes in *gimazjum*. When comparing the urban and rural areas, the importance of the school level ($\Delta Var_s$) increases by $15pp$ and the importance of the class level ($\Delta Var_c$ ) by $8pp$. Comparing this with the counter-factual difference in the control group,[13] provides an upper bound estimate. For the urban vs. rural, the change in sorting between is $\Delta Var_s^{URBAN} - \Delta Var_s^{RURAL} = 15pp - (-21pp) = 36pp$, whereas the

[12] That is, only Assumption 1 is satisfied.

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

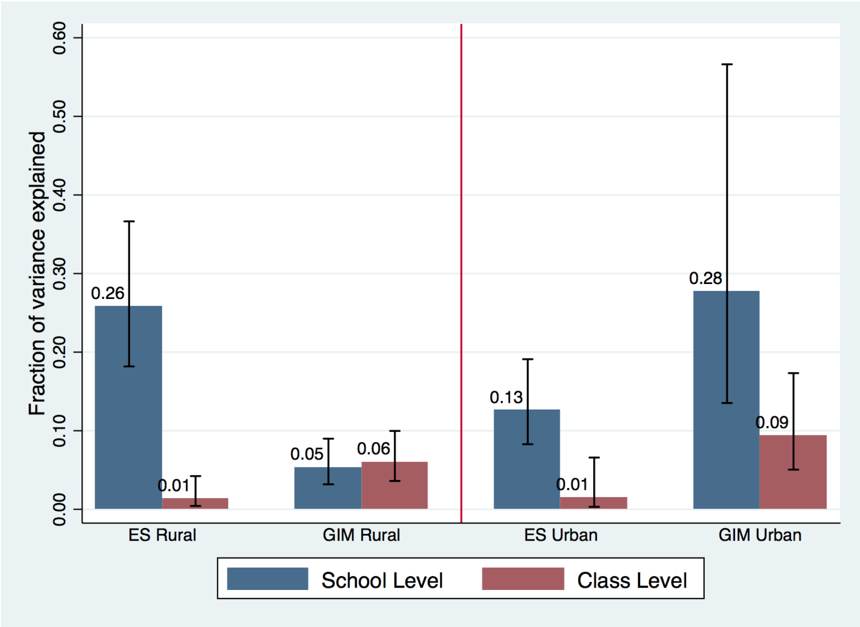


Figure 3: Decomposition of the Variance of Raven's Score - Urban vs. Rural

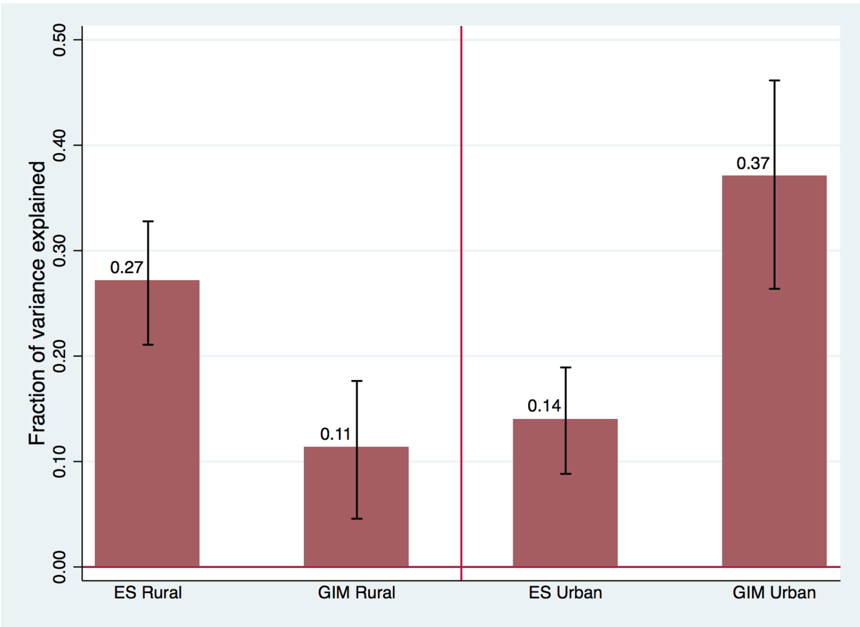


Figure 4: Residual fraction of the Variance of Raven's Score - Urban vs. Rural

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

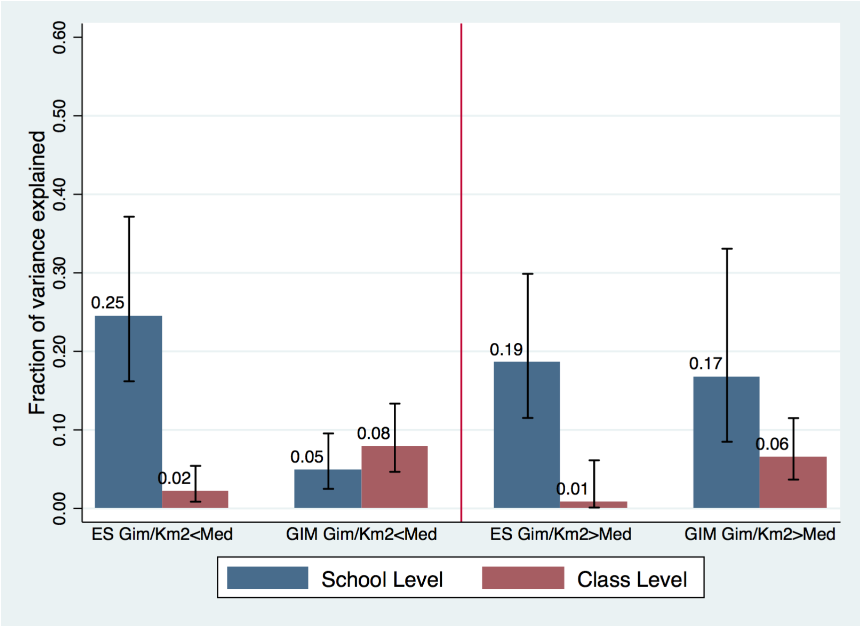


Figure 5: Decomposition of the Variance of Raven's Score - High vs. Low School Density

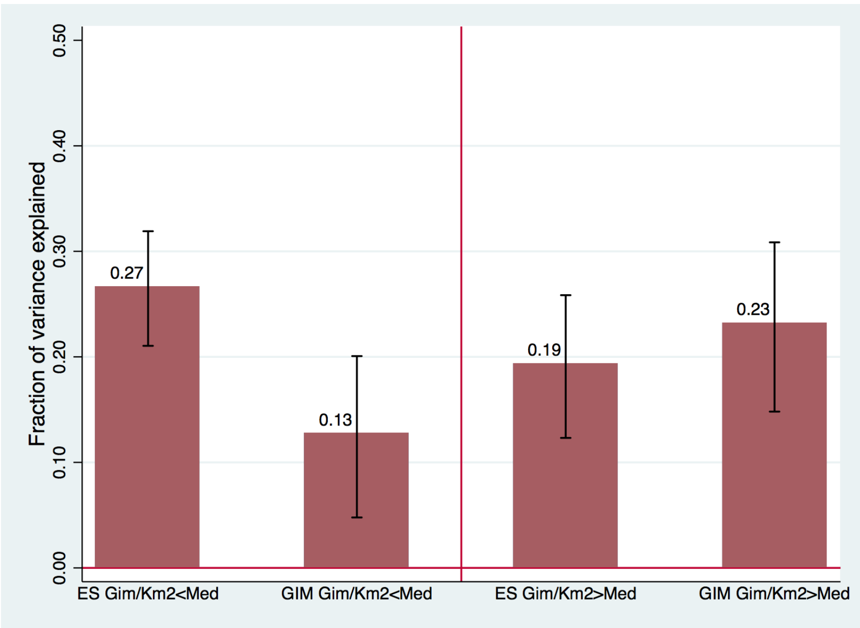


Figure 6: Residual fraction of the Variance of Raven's Score - High vs. Low School Density

Note: The figures present the decomposition of the variance of the standardized Raven's Progressive Matrix Score, by the school and class level, and the the unexplained (residual) proportion. The Estimation uses the mixed (hierarchical) effect model. High (low) $Gim/Km^2$ schools are in municipalities with the density of *gimnazja* per $km^2$ above (below) and above its median.

change in sorting within is $\Delta Var_c^{URBAN} - \Delta Var_c^{RURAL} = 8pp - (-16pp) = 24pp$. For the areas with high vs. low school density, the results are $18pp$ and $17pp$ respectively. Table 7 summarizes the calculations.

Table 7: The Effects of Interest

| Interpretation | **Lower Bound** | | **Upper Bound** |
|---|---|---|---|
| | *Urban* | *Rural* | *Difference* |
| Within | $9pp - 1pp = 8pp$ | $6pp - 1pp + (5pp - 26pp) = -16pp$ | $24pp$ |
| Between | $28pp - 13pp = 15pp$ | $5pp - 26pp = -21pp$ | $36pp$ |
| | *High* $Gim/Km^2$ | *Low* $Gim/Km^2$ | *Difference* |
| Within | $6pp - 1pp + (17pp - 19pp) = 3pp$ | $8pp - 2pp + (5pp - 25pp) = -14pp$ | $17pp$ |
| Between | $17pp - 19pp = -2pp$ | $5pp - 25pp = -20pp$ | $18pp$ |

*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.

Assumption 3.a says that the change in a general classroom assignment practice is the same in the areas with different cost 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 ($24pp$ or $17pp$ ). However, Assumption 3.b is unlikely to be true, and the mixing effect of larger catchment should be bigger in the areas with high cost of school choice. Replacing this assumption with that the mixing effect is proportional to the ratio of elementary schools to *gimnazja* might shed light on the possible magnitude of the actual effect. Table 1 shows that the ratio for the rural area is 2.31 elementary schools per *gimnazjum* and for the urban areas, the ratio is 1.49. From Table 5, in the rural areas sorting between schools drops by $21pp$ between the two stages of education. Hence, "back of the envelope" calculations suggest that the mixing effect in the urban areas is: $1.49/2.31 = 0.651$ times $21pp$, which equals $13.7pp$. Based on this, the effect of school competition on sorting between schools is $15pp + 13.7pp = 28.7pp$ of the proportion of variance explained by the school level. For the areas with low density of *gimnazja* the ratio is 3.12, and for the high density 1.81, consequently, the effect on sorting between is $-2pp + (1.81/3.12) * 20pp = 9.6pp$ of the proportion of variance explained by the school level. The larger effect of school competition in the case of urban vs. rural measure is not surprising given that this division is more contrasting than the measure based on school density (see Table 1).

### 5.3 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. The dog will affect a correlation between the test scores of students from the same test-room and thus a measure of their homogeneity. The problem is potentially serious, as students in *gimnazja* took Raven's test in groups, while in the elementary schools separately. This difference would imply more homogeneous classes in *gimnazja*. There are three reasons why this scenario is unlikely. Firstly, the team of professional psychometricians conducted the measurement with all the measures taken to provide a neutral environment for the test-takers (Jasinska, Hawrot, Humenny, Majkut and Konlewski, 2013). Secondly, the nature of these shocks would have to be different between the areas with different cost of school choice, and there are no reasons to suspect that. Thirdly, to fully exclude this possibility, I exploit the fact that in almost one-third of *gimnazja* students took Raven's test in two groups within a class. Thanks to this, I can directly check whether there is any impact of being in a separate group on Raven's score after controlling for the class fixed effects. The potential significant effect would indicate that the test-room environment matters for the outcome. The coefficients are, however, highly insignificant across the areas with different cost of school choice. On the other hand, the correlation between student's Raven's score and the average of 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 the rural areas, which is not consistent with the test-room shock story (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 the elementary school. 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 the 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 the areas with low cost of school choice, the difference is somehow higher, 4.8pp for the urban areas, and 4pp for the 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 the elementary school. Moreover, there was no institutional change which would have provided additional motivation for student grouping or mixing. Finally, the potential confounding effect would have to affect sorting differently across the areas with different cost of school choice. I find this possibility rather unlikely.

Table 8: Attendance at Local Elementary School

| Stage | All | Urban | Rural | High $Gim/Km^2$ | Low $Gim/Km^2$ |
|---|---|---|---|---|---|
| Elementary school | 79.1% | 72.4% | 82.1% | 76% | 81% |
| *Gimnazjum* | 82% | 77.2% | 83.4% | 80% | 84% |
| Difference | 2.9pp** | 4.8pp | 1.3pp | 4pp | 3pp |
| N | 10528 | 3455 | 7073 | 5218 | 5141 |

*Source*: the author's calculation based on the EVA survey. *Note*: Columns (Urban) and (Rural) show the statistics for the rural and urban schools, where the urban school are in municipalities with population larger than 50 000. Columns (Low $Gim/Km^2$ ) and (High $Gim/Km^2$ ) are for the areas with the density of *gimnazja* per $km^2$ below and above its median. Percentage of answers "yes" for the question asked to parents whether their child attended a local and assigned elementary school. *** denotes significance at the 1% level, ** at the 5% level.

## 6 Explaining Sorting Within a School

The results show that school competition leads to higher sorting of students between schools and within a school. While the former effect has been extensively investigated in the literature (Epple and Romano, 1998; Ladd and Fiske, 2000; Hsieh and Urquiola, 2006; Nechyba, 2006; Böhlmark et al., 2015), not much has been done to explore sorting across classes. This section briefly reviews the existing studies and explores the underlying mechanisms.

The empirical literature on determinants of sorting across classes is limited. Card and Rothstein (2007) show that the US schools, which 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 the US schools is far from random and might be detrimental for lower-achievers, as they also receive relatively fewer resources (e.g. novice teachers). Collins and Gan (2013) also report big variation in classroom assignment practices, but also show that grouping student together might be beneficial for low and high skill students, because of tailoring teaching practice. [14].

Two theoretical works directly link school competition with the sorting of students. Epple et al. (2002) argue that creation of a high track within a school might be used to attract high-skill or high-income students (the demand for peer quality), while Clotfelter et al. (2005) also suggest that it might be used to attract high-skilled teachers (the demand for teachers). Suppose that students differ by skill and maximise the expected difference between benefits and costs of education. The benefits are a function of peer quality and teacher skills, whereas the costs depend on a school distance. Students select a non-local school only if peers' and teachers' quality overbalance the extra costs of a longer travel distance. Next, suppose that school principals maximise enrollment and they 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 also 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 the low-quality area has to provide skilled teachers or a 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 mean 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).

The first part of this section presents the survey data on *gimnazjum*'s principals characteristics and their sorting practices. The second part empirically evaluates the two mechanisms mentioned above. The results suggest that the demand for peer quality is a likely explanation for the positive association between school competition and sorting within a school.

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

## 6.1 Survey Data on School Principals

The survey includes an open question about the class assignment [15] The principals underline the equal distribution of high and low achievers across classes, consistently across the areas with the different potential for the school competition. This practice stays in contrast with the finding of this paper. At that time there was a 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 *gimnazjum* in the rural or low school density areas.

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

## 6.2 The Demand for Peer Quality

School principals might create a high track within a school to attract non-local high-skill/income students or to keep the local ones. I correlate the *gimnazjum*-level measure of student sorting based on Raven's score with the sorting based on non-locality of students.[17] If a principal creates a high track to attract non-local students, there should be a positive association between the two measures of sorting. Since the focus is on the sorting at the entrance to *gimnazjum*, the observations from elementary schools are excluded. In particular, I follow Collins and Gan (2013) and define the sorting of students across classes as:

$$W_s^R = \frac{1}{2} \sum_c \frac{\widehat{\sigma}_{cs}^R}{\widehat{\sigma}_s^R} \tag{7}$$

[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]Unfortunatley, the available data do not allow to check whether school principals use sorting to keep local high-achievers.

Table 9: *Gimnazjum*'s Principals

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

Note: Urban and Rural show the statistics for the rural and urban schools, where the urban school are in municipalities with population larger than 50 000. Low $Gim/Km^2$ and High $Gim/Km^2$ are for the areas with the density of *gimnazja* per $km^2$ below and above its median. Variable "6th grade exam as a good signal" 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 the external examination allows 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 above variables equals one for questions:"strongly agree"/"rather agree" and 0 for "rather disagree"/"strongly disagree". Variable "Usage of the 6th grade exam" is one if principal's school analyzed examination score and used them somehow. *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.

## Table 10: The Demand for Peer Quality and The Demand for Teachers

| | All | | Urban | | Rural | | High $Gim/Km^2$ | | Low $Gim/Km^2$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| *Panel A: Dependent Variable: Sorting based on Raven's Score* ($W_s^R$) | | | | | | | | | | |
| Sorting based on Non-Locality ($W_s^N$) | .010 | | .254 | | -.055 | | .052 | | -.030 | |
| | (.054) | | (.070)*** | | (.055) | | (.071) | | (.076) | |
| $\lvert NonLocal_{1s} - NonLocal_{2s} \rvert$ | | -.096 | | -.209 | | .021 | | -.128 | | -.030 |
| | | (.046)** | | (.052)*** | | (.047) | | (.048)*** | | (.109) |
| Constant | .948 | .970 | .699 | .987 | 1.013 | .959 | .912 | .983 | .983 | .957 |
| | (.050)*** | (.008)*** | (.068)*** | (.014)*** | (.050)*** | (.009)*** | (.066)*** | (.009)*** | (.070)*** | (.015)*** |
| N | 136 | 141 | 41 | 44 | 95 | 97 | 69 | 72 | 67 | 69 |
| $R^2$ | .0005 | .043 | .214 | .268 | .019 | .002 | .012 | .119 | .005 | .002 |
| *Panel B Dependent Variable: Mean of Raven's score at the class level* | | | | | | | | | | |
| Teacher's Rank | -.080 | -.096 | -.025 | -.223 | -.100 | -.081 | .099 | .028 | -.287 | -.239 |
| | (.080) | (.101) | (.173) | (.222) | (.087) | (.111) | (.093) | (.146) | (.101)*** | (.126)* |
| Teacher's Experience | | .002 | | .026 | | -.003 | | .009 | | -.007 |
| | | (.008) | | (.019) | | (.009) | | (.012) | | (.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 results of the OLS regression of the standardized Raven's score sorting measure $W_s^R$ on the non-locality sorting measure $W_s^R$ 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 (*gimnazjum*). Panel B presents the OLS regressions of the class average of the standardized Raven's score on class-averages of teacher experience in years and teacher's professional rank. The rank ranges from 0 to 5, where 0 is rank-less 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*). 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.

where $\widehat{\sigma}_{cs}^R$ is the observed standard deviation of Raven's score for class $c$ from *gimnazjum* $s$ and $\widehat{\sigma}_s^R$ is the observed standard deviation of Raven's score for *gimnazjum* $s$. The ratio is at the 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 is positive. Hence the measure $W_s^R$ is null. With perfect mixing, the variance at the class-level is the same as at the school-level and $W_s^R$ is one.[18] I define a similar measure for the sorting based on non-locality of students:

$$W_s^N = \frac{1}{2} \sum_c \frac{\widehat{\sigma}_{cs}^N}{\widehat{\sigma}_s^N} \tag{8}$$

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

$$W_s^R = \alpha + \beta W_s^N + \epsilon_s \tag{9}$$

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

The measure of sorting based on non-locality of students might be misleading when there are a few non-local students. Suppose that there is one non-local student and she is randomly assigned to the class. For the assigned class the measure will be one, and for the second class, it will be zero. Consequently, the school-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 nonlocality. Since there are only 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:

$$W_s^R = \alpha + \beta |NonLocal_{1s} - NonLocal_{2s}| + \epsilon_s \tag{10}$$

[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.

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

### 6.3 The Demand for Teachers

High-quality teachers can improve the attractiveness of a school. Because the teacher wages are fixed, the principals might create homogeneous and high-track classes to attract skilled teachers, assuming they prefer such environment. To test for this possibility, I correlate teacher characteristics with the class average of Raven's score and control for the school fixed effects. A positive association between the measures of teacher experience and the class-average of Raven's score is consistent with the demand for teachers hypothesis. The focus is only on teachers and classes from *gimnazja*. The regression of interest is specified as follows:

$$\overline{Y}_{cs} = \alpha + \beta T_{cs} + \mu_s + \epsilon_{cs} \tag{11}$$

where $\overline{Y}_{cs}$ is the average Raven's Score for class $c$ from *gimazjum* $s$, $T_{cs}$ is the class-average of teacher characteristics and $\mu_s$ are the school fixed effects. I use two measures of teacher characteristics: teaching experience in years and the professional rank. The rank ranges from the rankless teacher (=0) to "the professor of education" (=5).

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

The results suggest that *gimnazjum* principals do not offer high-tracks to attract skilled teachers. Interestingly, the results, even though imprecise, suggest that school principals might assign higher-rank teachers to worse classes in the urban areas and the areas with low school density. They might want to compensate lower-peer quality or increase discipline with better teachers.[19] Regardless of the reasons, this might have a positive effect on educational equality of opportunity. However, more data is needed to investigate this possibility fully.

# 7 Conclusions

This study shows that school competition leads to a higher sorting of students within a school and between schools. It links the sorting across classes with student's demand for peer quality, which motivates school principals to create selective tracks. The results bear relevance for policy makers who wish to use school competition as a mean 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. The 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 also for low-background or low-performing students (MacLeod and Urquiola, 2009). Even though the value-added based accountability has been heavily discussed, not much attention has been paid to the potential distributional effects (Rothstein, 2009; Angrist, Pathak and Walters, 2011; Chetty, Friedman and Rockoff, 2014). The alternative policy could be to link school vouchers with the socioeconomic background, for instance, to offer them only to students with low-income [20]. On the other hand, abolishing the teacher collective bargaining agreements allows school principals to compete based on wages rather than a composition of students.

# References

**Angrist, Joshua D, Parag A Pathak, and Christopher R Walters**, "Explaining charter school effectiveness," Technical Report, National Bureau of Economic Research

[19] An alternative explanation is that high-skill teachers prefer to teach low-skill students. However, in this scenario, high-skilled students will be not 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).

2011.

**Betts, Julian R and Jamie L Shkolnik**, "Key difficulties in identifying the effects of ability grouping on student achievement," *Economics of Education Review*, 2000, *19* (1), 21–26.

**Böhlmark, Anders, Helena Holmlund, Mikael Lindahl et al.**, "School choice and segregation: evidence from Sweden," Technical Report 2015.

**Brouwers, Symen A, Fons JR Van de Vijver, and Dianne A Van Hemert**, "Variation in Raven's Progressive Matrices scores across time and place," *Learning and Individual Differences*, 2009, *19* (3), 330–338.

**Burris, CC and KE Allison**, "Review of "Does sorting students improve test scores?"," *Boulder, CO: National Education Policy Center. Retrieved February*, 2013, *11*, 2014.

**Card, David and Jesse Rothstein**, "Racial segregation and the black–white test score gap," *Journal of Public Economics*, 2007, *91* (11), 2158–2184.

**Chetty, Raj, John N Friedman, and Jonah E Rockoff**, "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *The American Economic Review*, 2014, *104* (9), 2593–2632.

**Clotfelter, Charles T, Helen F Ladd, and Jacob Vigdor**, "Who teaches whom? Race and the distribution of novice teachers," *Economics of Education review*, 2005, *24* (4), 377–392.

**Collins, Courtney A and Li Gan**, "Does sorting students improve scores? An analysis of class composition," Technical Report, National Bureau of Economic Research 2013.

**Dolata, Roman**, *Szkoła, segregacje, nierówności*, Wydawnictwa Uniwersytetu Warszawskiego, 2008.

**Duflo, Esther, Pascaline Dupas, and Michael Kremera**, "Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya," *The American Economic Review*, 2011, *101* (5), 1739–1774.

**Epple, Dennis and Richard E Romano**, "Competition between private and public schools, vouchers, and peer-group effects," *American Economic Review*, 1998, pp. 33–62.

_ **, Elizabeth Newlon, and Richard Romano**, "Ability tracking, school competition, and the distribution of educational benefits," *Journal of Public Economics*, 2002, *83* (1), 1–48.

**Figlio, David N and Marianne E Page**, "School choice and the distributional effects of ability tracking: does separation increase inequality?," *Journal of Urban Economics*, 2002, *51* (3), 497–514.

_ **and Maurice E Lucas**, "What's in a grade? School report cards and the housing market," *American Economic Review*, 2004, pp. 591–604.

**Friedman, Milton**, *The role of government in education*, Rutgers University Press, 1955.

**Gibbons, Stephen, Stephen Machin, and Olmo Silva**, "Choice, competition, and pupil achievement," *Journal of the European Economic Association*, 2008, *6* (4), 912–947.

**Herbst, Mikołaj, Jan Herczyński, and Anthony Levitas**, *Finansowanie oświaty w Polsce: diagnoza, dylematy, możliwości*, Wydawnictwo Naukowe Scholar, 2009.

**Hoxby, Caroline M**, "Does competition among public schools benefit students and taxpayers?," *American Economic Review*, 2000, pp. 1209–1238.

**Hsieh, Chang-Tai and Miguel Urquiola**, "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program," *Journal of Public Economics*, 2006, *90* (8), 1477–1503.

**Instytut Badań Edukacyjnych**, "Raport o stanie edukacji 2010," 2011.

**Jasinska, Aleksandra, Joanna Hawrot, Grzegorz Humenny, Przemyslaw Majkut, and Maciej Konlewski**, "Kontekstowy model efektywnosci nauczania po I etapie edukacyjnym - wyniki badania podluznego w szkolach podstawowych.," Technical Report, Instytut Badan Edukacyjnych 2013.

**Jensen, Arthur Robert**, *The G Factor: The Science of Mental Ability*, Praeger, 1998.

**Kalogrides, Demetra, Susanna Loeb, and Tara Béteille**, "Systematic sorting teacher characteristics and class assignments," *Sociology of Education*, 2013, *86* (2), 103–123.

**Kane, Thomas J, Stephanie K Riegg, and Douglas O Staiger**, "School quality, neighborhoods, and housing prices," *American Law and Economics Review*, 2006, *8* (2), 183–212.

**Kern, Nora, Renita Thukral, and Todd Ziebarth**, "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*, 2012.

**Kremer, Michael, Esther Duflo, and Pascaline Dupas**, "Peer Effects, Teacher Incentives, and the Impact of Tracking," *American Economic Review*, 2011, *101* (5 (August 2011)).

**Ladd, Helen and Edward Fiske**, "When schools compete: A cautionary tale," *Washington, DC: The Brookings Institution*, 2000.

**MacLeod, W Bentley and Miguel Urquiola**, "Anti-lemons: school reputation and educational quality," Technical Report, National Bureau of Economic Research 2009.

**Meghir, Costas and Mårten Palme**, "Educational reform, ability, and family background," *The American Economic Review*, 2005, *95* (1), 414–424.

**Nechyba, Thomas J**, "Income and peer quality sorting in public and private schools," *Handbook of the Economics of Education*, 2006, *2*, 1327–1368.

**Raven, John**, "The Raven's progressive matrices: change and stability over culture and time," *Cognitive Psychology*, 2000, *41* (1), 1–48.

**Rothstein, Jesse**, "Student sorting and bias in value-added estimation: Selection on observables and unobservables," *Education*, 2009, *4* (4), 537–571.

**Szmigel, Maria Krystyna**, "Gimnazjum niejedno ma imię? - na podstawie wywiadoów z dyrektorami. Doniesienie z badań jakościowych," Technical Report, XIX Konferencja Diagnostyki Edukacyjnej, Gniezno 2013 2013.

**Tiebout, Charles M**, "A pure theory of local expenditures," *The Journal of Political Economy*, 1956, pp. 416–424.