Hidden Tracking or Mixing? The Educational Market and Sorting of Students within the Polish Comprehensive Education.

Paweł Bukowski*

In order to reduce educational inequalities, Poland in 1999 expanded and changed the structure of the comprehensive education. Two separate comprehensive schools have been created at the place of the single one. This paper investigates the relationship between the type of educational market, its competitiveness and the amount of sorting within the expanded comprehensive education in Poland. On the one hand, at the transition between comprehensive schools in more competitive and denser educational markets, students might self-select into schools and principals may strategically respond to this. The expanded comprehensive education might be then *de facto* a combination of a comprehensive and tracking education. On the other, in less competitive markets it might be truly comprehensive and further overcome initial sorting. The empirical results confirm this intuition. Compare to the initial sorting at the first comprehensive school, the secondary comprehensive school in the urban areas is in reality a form of hidden tracking, while in the rural it is effectively mixing students. Moreover, I find evidence that higher school competition is connected with higher sorting across classes and between schools within the comprehensive education.

*Department of Economics, Central European University, Budapest, Hungary; e-mail:bukowski_pawel@ceu-budapest.edu

1 Introduction

The length of comprehensive education, relative to tracking[1], is at the core of the debate about educational inequalities (Betts 2011).[2] In the European context[3], the literature shows that expansion of comprehensive education (and shortening tracking) reduce income inequalities (Meghir and Palme 2005), inter-generational income correlation (Brunello and Checchi 2007, Pekkarinen et al. 2009) and dispersion of student achievement measures (Gamoran 1996, Ammermüller 2005, Hanushek and Woessmann 2007, Horn 2009).[4] Because of these reasons, the reforms aimed at reducing tracking became main policy tools to fight educational inequalities of opportunities (in i.e. Sweden, Finland or Norway). The Polish education system experienced similar changes in 1999. Because it has been always characterized by relatively strong educational inequalities of opportunities (Bukowski and Kobus 2012), the main component of the reform was the expansion of the comprehensive (and compulsory) education from 8 to 9 years. However, it has been also split into two separate schools: a 6-year elementary school and a 3-year lower secondary school (called *gimnazjum*). The admission to these schools is based on catchment areas, however parents may request an alternative school. In addition to this, standardized examinations have been introduced after each stage of education.

Even though the first 9 years of education in Poland is *de jure* comprehensive, in some local educational markets it might be *de facto* a combination of a comprehensive and tracking education. This is because at the entrance to *gimnazjum* students are relatively more mobile, their achievements are known and there is a clear ranking of these schools. As a result, *gimnazja* face more intense competition than elementary schools, which may lead to a strategic behaviour of school principals and create additional stage of sorting

[1] In the comprehensive education every student has a right to attend a local school and admission process is not based on any measure of previous achievement. Conversely, in the tracking education, students are sorted into schools or classes based on their merit.

[2] By this term I mean both: the dispersion of a distribution of a measure of student performance and the educational inequalities of opportunities. When it matters, I make this distinction explicit in the text.

[3] In Europe students are usually sorted into different schools, while in the US - into different classes within a school. Importantly, in many countries (including Poland) students are assigned into a specific class, which is not changed across courses. Therefore the composition of peers is often fixed during each stage of education.

[4] Nevertheless, there are few studies which show that there is no negative effect of tracking (Galindo-Rueda and Vignoles 2005, Waldinger 2006). In the US, the early studies show a negative effect of tracking on inequalities (Hoffer 1992, Argys et al. 1996), however they may fail to establish the causal relationships. The later works, which try to alleviate the endogeneity problems, find null (Betts and Shkolnik 2000a, Zimmer 2003) or decreasing (Figlio and Page 2002) effect of tracking on inequalities.

within the comprehensive education. On the other hand, in less competitive markets the expanded comprehensive education might be truly comprehensive and the secondary comprehensive school might further overcome initial sorting. This is because there is less *gimnazja* than elementary schools, which leads to further mixing of students.

This paper investigates the relationship between the type of educational market, its competitiveness and the amount of sorting within the expanded comprehensive education in Poland. I consider sorting at the entrance to elementary school as a reflection of residential sorting and general student mobility specific for a given educational market. Next, I'm looking on how different educational markets influence sorting into *gimnazjum*, *relative* to the initial sorting.[5]

I find evidence for the heterogeneous effect. At the transition between the corresponding stages of the comprehensive education, in the rural areas sorting within a school and between schools becomes weaker (students are more heterogeneous at the class and school levels), while in the case of the urban areas both types of sorting are reinforced. In other words, compare to the initial sorting at the elementary school level, *gimnazjum* in the urban areas is in reality a form of hidden tracking, while in the rural it is effectively mixing students. Moreover, schools with higher share of non-local students are on average more homogeneous and have higher sorting across classes.

Firstly, the school network and competition are likely reasons for the differential change in sorting between schools. Because of a school choice, higher parental education and low commuting costs, in the urban areas parents and students are more likely to select alternative comprehensive school than in the rural areas. At the same time, because of the clear economies of scale, school principals have motivation to compete between each other for students. According to the school choice literature (Tiebout 1956, Epple and Romano 1998, Hoxby 2000, Rothstein 2007, Hsieh and Urquiola 2006) the more competitive educational market should generally lead to a higher sorting between schools. On the other hand, the fact that there are more elementary schools than *gimnazja* is directly leading to mixing of students between schools. This is true in both areas, but is likely to be the only force in the rural areas.

Secondly, I argue that the heterogeneous change in sorting within a school might be a result of competition. Theoretical model developed by Epple et al. (2002) shows that, as long as parents consider peer group quality, school principals might want to

[5] This approach is somehow similar to the difference-in-difference methodology, where change in mechanisms on the specific educational market ("treatment" - i.e. higher school competition caused by a school ranking interacted with high student mobility) influences flow of students from the first to the second stage of comprehensive education ("before" and "after").

attract them by creation of a high track within a school. Despite the political pressure to randomize class composition,[6] I observe increase in sorting within urban *gimnazja*. Anecdotal evidence suggests that language classes might serve as a signaling device, for instance: a French language class effectively excludes the low-class students from the peer group. Contrary to this, the lack of competition leads to mixing of rural students across classes.[7]

The focus of this paper is on the interplay between characteristics of the educational market and sorting of student, however it would be also informative to estimate a similar effect of the change in the structure of the comprehensive education. The 1999 educational reform has replaced the single 8-year long elementary school by two schools: the new 6-years elementary school and 3-years *gimnazjum*. I argue that my research design can provide me with the effect of the change in the regime, which is biased toward higher sorting. This means that in the case of rural areas I estimate the lower bound of the effect and the actual increase in mixing of student was at least as large as the reported change. Contrary to this, in the case of the urban areas I estimate the upper bound, so in reality the change in sorting might be neutral. Nevertheless, the result clearly show that splitting comprehensive education is potentially a very effective tool in fighting inequalities of opportunities in areas with low competition between schools.

The contribution of this paper is threefold. Firstly, it underlines the importance of the local educational market competitiveness for educational inequalities. Understanding the mechanisms, which lead to student sorting will allow to design policy tools, which fit better to local conditions. Secondly, the positive interplay between a competition and sorting within a school contributes to the school choice literature, which mostly focuses on sorting between schools. This suggests, that sorting practices at the school level might be a strategic decision of school principals. Moreover, the attention so far has focused on the effect of the private educational sector. In this research I show that similar effects may appear in the markets with marginal role of the private sector. Finally, this research argues that the trade-off between length of tracking and comprehensive education is just part of the picture. The structure of the later might introduce additional stages of sorting or mixing and thus be not neutral for inequalities. Therefore, the attention of scholars and policymakers should not only focus on *for how long* to keep students in the

[6]In Poland, within a given stage of education assignment of students into classes is fixed across grades and subjects.

[7]The political pressure toward reducing inequalities might be a reason behind this practice.

comprehensive part of education, but also more generally *how*.[8]

The paper is organized as follows. In the section 2 I discuss sorting in the Polish education system. The section 3 is devoted to the research design and data. In the section 4 I show the results and robustness checks. In the section 5 I interpret the results and finally, in the section 6, I conclude.

2 Institutional Background and Sorting of Students

The Polish comprehensive and compulsory education system consists of 6 years of elementary school, which is then followed by 3 years of *gimnazjum*. The admission to the comprehensive schools is based on catchment areas, which means that every student living within this area has a right to attend a given public school. However, parents may request an alternative school, but its principal has a right to reject the application. The elementary school and *gimnazjum* are completely separated entities, with different managerial and teaching body, but usually they serve the same community of students. Because there are more elementary schools than *gimnazja*[9], the catchment area for the later is usually larger and contains the catchment areas from local elementary schools. During the comprehensive education, students are examined by the two standardized, externally graded and obligatory examinations: a low stake after elementary school (6th grade) and a high stake after *gimnazjum* (9th grade). The later serves as a basis for the admission into the higher secondary education, which is a first part of tracking. Students can choose a track (academic, mixed or vocational) and apply to any high schools, but the admission is not granted.

At the entrance to each school students are assigned into classes. Unlike the other educational systems (i.e. the US) this assignment is fixed across grades and subjects. The reallocation across classes are allowed only in special situations, therefore usually the peer composition of classes is constant within each stage of education.

[8] Not much has been done to analyze this in other settings. For Poland, Dolata (2011) shows that introduction of *gimnazjum* reinforces sorting between urban schools. However, this work suffers from the identification issues - it's hard to establish causal effect as the result might reflect sorting at the elementary school or common education process. In my work, I alleviate these problems by the comparison of background (pre-determined) characteristics between students at both elementary school and *gimnazjum*. Thanks to this, I can separate sorting of students at the entrance to both stages of comprehensive education.

[9] Most of elementary schools have been constructed during the past 50 years, while *gimnazja* have appeared just decade ago. The network of elementary schools thus reflects the past demographic situation and it is considered as too dense. The network of *gimnazja*, in turn, is more "rational" in the sense that it is better adjusted to the current demographic needs. In addition to this, elementary schools serve younger children for whom distance to a school matter more than for older children.

2.1 Sorting Between Schools

The first and the most basic reason why students are similar to each other within a school is residential sorting. Because of various reasons (i.e. neighborhood quality, local economic conditions or historical accidents) similar people tend to live together within a school catchment area (Tiebout 1956). At the same time, the housing prices might be influenced by a local school characteristics, which further reinforces self selection (Figlio and Lucas 2004, Kane et al. 2006). As a result, residential sorting is often the main source for differences in student composition across schools.

The second source is a school choice. Lower financial and informational constraints may lead high-class parents to send their kids to an alternative school with a better reputation. Therefore, as shown in Epple and Romano (1998), the larger and more competitive educational market should generally lead to higher sorting. Nevertheless, in the case of rural areas the educational market is limited - the school choice is small and the potential cost of sending a child to a non-local school is relatively high.[10] This is not the case in urban areas, where transportation costs are low, number of schools higher and parents more educated.

Table 1: The School Network in Poland 2012/13

Stage	All	Urban	Rural
Elementary Schools	12 696	4092	8604
Gimnazja	6470	2748	3722
Ratio	1.96	1.49	2.31

source: Herczynski and Sobotka (2013, p.38,49-50). Ratio is the number of elementary schools divided by the number of *gimnazja*.

How *gimnazja* changes intensity of sorting between schools is ambiguous. Firstly, there are less *gimnazja* than elementary schools and since both stages are obligatory, students from different elementary schools will be mixed together in one *gimnazjum*. This decreases sorting between schools. Table 1 shows ratio of the number of elementary schools to *gimnazja* in a rural-urban breakdown. There are on average around two elementary school per one *gimnazjum*, but this number differs between the urban and rural areas with the former having much smaller ratio. Even though the mixing effect

[10] It includes a transportation cost, missing links with peers from the neighborhood, limited possibilities of grass-root actions with other parents.

should take place in both areas, it will be more visible in the rural schools. On the other hand, since elementary schools are separated from *gimnazja* and older students are more mobile, parents may want to adjust their choice and send children to *gimnazjum* other than a local one. Thanks to the standardized examinations parents have an access to easy-accessible rankings of *gimnazja*, which limit the informational constraints and thus makes the selection of school easier. The sorting effect should be more important in the urban areas.

2.2 Sorting Within a School

In contrast to sorting between schools, sorting within a school is mainly determined by a school's principal (who, in turn, might be influenced by parents or other agents). There are reasons why she might want to sort students across classes. Firstly, there might be an advantage for principals and parents for placing students from one neighborhood into one class. Parents can share cost of transportation, help each other in the case of school problems (like conflict with a teacher) and principals can coordinate a course schedule with a school transportation. Nevertheless, because of residential sorting, this will result in grouping similar students together in one class. Another reason might be cost of extra school activities, which are shared by parents i.e. school excursions. Grouping poor and rich students separately allows to adjust school activities to the parental budget. Finally, adjustment of class composition based on language knowledge might also reproduce the class division of the society. Since more advantaged students are more likely to receive language classes before entering a school they will be placed into one class.[11]

Gimnazjum should generally lead to an increase in sorting within a school when schools are competing for students or students are coming from diverse environments. Firstly, the standardized examination after the sixth grade (just before entering *gimnazjum*) provides principals with relatively accurate signal about students' abilities (they are generally unknown at the entrance to elementary school). Therefore it is relatively easy for principals of *gimnazja* to create specialized classes, which may attract parents and students. Secondly, similarly as at the entrance to elementary school, principals may want to sort students based on their language knowledge (or other specific abilities, i.e. sport). After the sixth grade however, the difference between students are likely to be larger, and as long as it's more connected with students' background, the homogeneity of classes may increase. On the other hand, there is a political pressure to decrease

[11] Perhaps high-class parents know about this and may strategically send their kids to an extra language course

educational inequalities, therefore principals may want to randomize class composition.[12]

3 Empirical Strategy and Data

3.1 Data

The data are drawn from the first wave (2010) from the panel of the sample of Polish students created by the Educational Value Added Team.[13] The cross-section consists of almost 6000 first-graders and 6000 sixth-graders (which is the first grade of *gimnazjum*) from 360 randomly drawn schools in Poland. The main outcome variable and measure of background characteristics is a standardized (separately for the first and sixth graders) cumulative score from the Raven's Progressive Matrix test. For each student's i from grade g, I calculate so-called a z-score:

$$zscore_{ig} = \frac{score_{ig} - sc\bar{o}re_g}{sdscore_g}$$

where $score_{ig}$ is a raw Raven score and $sdscore_{ig}$ is a standard deviation of Raven score for each grade. Beside this, the set of student, parental, school and county[14] characteristics are available. Importantly, it includes questions about each school's sorting practices. All the statistics used in the paper are weighted using an appropriate weighting scheme, thus the results should be interpreted as representative for the corresponding Polish populations. Table 2 summarizes the available sample.

[12]In the 5th Section I discuss a qualitative study of principals' policies toward a class composition. Majority of answers were clearly indicating willingness to make hetereogeneous classes, however they also underlined the need for sorting based on i.e. language knowledge.

[13]The Educational Value Added research team is a part of the Central Examination Commission, a Polish institution which is conducting obligatory exams for all students in Poland

[14]Because of the school anonymity reasons I don't have access to the municipality level data.

Table 2: *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 sample										
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 sample										
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				

Note: The descriptive statistics are calculated for the sample, not for the population, therefore no weighting is used.

The Raven's Progressive Matrices test, developed by John C. Raven in 1936, is the most popular test, which is aimed at the general intelligence. The test usually consists of 4x4 3x3 or 2x2 matrix of figures at each entries (except the lowest diagonal which is empty). Figures in each row are following the same pattern and the task of the subject is to identify the missing element according to this pattern. It is designed to capture two abilities: eductive ability (to make a meaning from confusion) and reproductive ability (to absorb, recall and reproduce explicit information) (Raven 2000). This test has been used in "Cross-cultural comparisons...[which] are often conducted from the premise that the instrument measures cross cultural differences in intelligence that are not confounded by other cultural or national differences, such as education and affluence. 'Culture-free'..., 'culture-fair'..., and 'culture-reduced'.. are all terms that have been proposed to describe the Raven or similar tests that do not seem to require much cultural knowledge for answering the items correctly." (Brouwers et al. 2009). The Raven's test depends on biological characteristics which are determined by the genotype and parent's behavior during the pregnancy and early childhood. Therefore, used in the measurement of inequalities, is not a source of the reflection (simultaneity) problem.[15]

3.2 Empirical Strategy

In order to see how general level of school homogeneity changes across two comprehensive schools and educational markets I apply an empirical approach similar as in the peer effect literature (Sacerdote 2001). I run the following regression of a standardized student's own Raven's score on the leave-out mean of schoolmates, separately for the rural and urban areas:

$$Y_{ics} = c + \alpha_1 \overline{Y}_{-is} + \alpha_2 GIM_s + \alpha_3 \overline{Y}_{-is} \times GIM_s + \epsilon_{ics} \tag{1}$$

Where Y_{ics} denotes the outcome for student i from class c in school s. $\overline{Y}_{-is}$ is the school-level leave-out mean, GIM_s denotes observations coming from *gimnazjum*. α_3 - captures the impact of *gimnazjum* on the level of homogeneity. Please note that I'm not trying to capture any causal relation between a student's and her peers' characteristics. These are pre-determined and the only source of the correlations is sorting.

To separate sorting between schools and sorting within a school I run two additional regressions with different independent variables: the leave-out mean of Raven's score of

[15] The reflection problem emerges when an individual's outcome is influenced by her peers' outcome, but at the same time their outcome is also influenced by her. This causes the endogeneity problem and bias estimates. See Manski (1993), Sacerdote (2001).

classmates $\overline{Y}_{-ics}$ (2) and the mean of Raven's score of outside-class schoolmates $\overline{Y}_{-cs}$ (3). As previously, the regressions are also estimated separately for the rural and urban areas.

$$Y_{ics} = c + \beta_1 \overline{Y}_{-ics} + \beta_2 GIM_s + \beta_3 \overline{Y}_{-ics} \times GIM_s + \epsilon_{ics} \tag{2}$$

$$Y_{ics} = c + \gamma_1 \overline{Y}_{-cs} + \gamma_2 GIM_s + \gamma_3 \overline{Y}_{-cs} \times GIM_s + \epsilon_{ics} \tag{3}$$

Similarly to the previous equation β_3 and γ_3 capture the impact of *gimnazjum*. Please note that since some schools in the sample have only one class per grade, the number of observations for the regression (3) is smaller.

The identification of sorting within and between is based on comparison of the coefficients from the aforementioned regressions. Both types of sorting increases the correlation of a student's own Raven's score with the leave-out mean of classmates. However, the correlation with outside-class schoolmates is positively affected by sorting between, but negatively by sorting within. Therefore in the presence of both types of sorting the correlation with classmates should be significantly higher than with outside-class schoolmates. The same logic applies to changes caused by *gimnazja*, Table 3 shows the expected sign of the coefficients β_3 and γ_3 for the combination of changes in sorting.

Table 3: The sign of parameters and increase in sorting

Incerease in Sorting:	Between	Within	Both
Corr. with classmates: β_3	0	+	+
Corr. with outside-class schoolmates: γ_3	+	-	?

The first part compares changes in sorting in the rural and urban areas. Nevertheless, this breakdown might be too vague to claim anything about the role of competitiveness. As an alternative specification, I use a measure of the competitiveness of the local educational market and see how it interacts with the change in sorting after the elementary school. Specifically, my proxy is a share of non-local students within a *gimnazjum*, which reflects the extent of a school choice in a local area. I define $nonlocal_s$ as a number of non-local first-graders and $total_s$ as a total number of first-grades in *gimnazjum* s, then my measure of competition is:

$$comp_s = \frac{nonlocal_s}{total_s}$$

It is very likely that schools with high number of non-local students are also located in the areas with denser network of *gimnazja* relative to elementary schools. Therefore the triple interaction of $comp_s$, the *gimnazjum* dummy GIM_s and the means $\overline{Y}_{-ics}$ or $\overline{Y}_{-cs}$ might capture the effect of the school network instead of competition. In order to control for this I also include the ratio of the number of elementary school to *gimnazja*: $ratio_s$ (presented in Table 1) in a county where school s is located. The expanded version of regressions (2) and (3) are thus:

$$\begin{aligned} Y_{ics} = c &+ \beta_1 \overline{Y}_{-ics} + \beta_2 GIM_s + \beta_3 \overline{Y}_{-ics} \times GIM_s \\ &+ \beta_4 \overline{Y}_{-ics} \times GIM_s \times comp_s \\ &+ \beta_5 \overline{Y}_{-ics} \times GIM_s \times ratio_s + \epsilon_{ics} \end{aligned} \tag{4}$$

$$\begin{aligned} Y_{ics} = c &+ \gamma_1 \overline{Y}_{-cs} + \gamma_2 GIM_s + \gamma_3 \overline{Y}_{-cs} \times GIM_s \\ &+ \gamma_4 \overline{Y}_{-cs} \times GIM_s \times comp_s \\ &+ \gamma_5 \overline{Y}_{-cs} \times GIM_s \times ratio_s + \epsilon_{ics} \end{aligned} \tag{5}$$

The coefficient β_4 (β_5) captures the relationship between the competitiveness (the school network) and the change in correlation with classmates in *gimnazjum*. Similarly γ_4 (γ_5) captures the relationship with the change in correlation with outside-class peers. In other words, together β_4 and γ_4 (β_5 and γ_5) show the relationship between competitiveness of the educational market (the school network) and change in sorting between and within at the transition between the two comprehensive schools.

4 Results

4.1 Results in the Urban and Rural Areas

The regressions results, which are presented in Table 4 should be interpreted in a more qualitative way. The sign of the coefficients reflects direction of changes in sorting, but as there is no natural scale, the magnitude itself is hard to interpret. In the following discussion I do not focus on exact values of the estimated parameters, rather I discuss the general effect on sorting and relative change in the correlations.

Table 4 Panel A presents results from the first regression (1). In the rural areas[16] elementary schools are more homogenous than *gimnazja*, as the correlation drops by almost 25%. Conversely, in the urban areas (if anything) *gimnazja* are more heterogeneous. The coefficient on the interaction terms shows an increase by 7%, however it is insignificant. Table 5 presents the Gini Coefficient of the distribution of the leave-out mean of Raven's score at the school level for combinations of the location and type of school. Consistently with the regression results, the peer quality is the most equal among the rural *gimnazja* and the urban elementary schools, which are followed by the rural elementary schools and the urban *gimnazja*.

Table 4: Regressions Result

	All	Urban	Rural
Panel A : 1st Regression: Leave-out mean at the school level			
$\overline{Y}_{-is}$	.872*** (.15)	.802*** (.032)	.870*** (.022)
$\overline{Y}_{-is} \times GIM_s$	-.091** (.03)	.059 (.037)	-.208** (.64)
GIM_s	-.002 (.008)	-.033** (.012)	-.005 (.013)
n	10496	3627	6869
Panel B : 2nd Regression: Leave-out mean at the class level			
$\overline{Y}_{-ics}$	.819*** (.02)	.694*** (.043)	.830*** (.028)
$\overline{Y}_{-ics} \times GIM_s$	-.074* (.031)	.16*** (.047)	-.196*** (.05)
GIM_s	.0005 (0.009)	-.049** (.16)	-.001 (.014)
n	10496	3627	6869
Panel C : 3rd Regression: Outside - class mean			
$\overline{Y}_{-cs}$	.714*** (.054)	.554*** (.088)	.750*** (.073)
$\overline{Y}_{-cs} \times GIM_s$	-.438*** (.093)	-.148 (.137)	-.633*** (.14)
GIM_s	-.017 (.021)	-.024 (.04)	-.035 (.031)
n	9768	3568	6044

Note: Robust standard errors in parentheses.**** denotes significance at the 0.1% level, ** at 1% level and * at the 5% level. The outcome variable is a standardized Raven's Progressive Matrix Test score (with subtracted mean and divided by standard deviation).

[16]An urban area consists of towns and cities above 50 thousand inhabitants. A rural area consists of villages and towns below 50 thousand inhabitants

Table 5: The Gini Coefficients of Peers Quality distribution

School & Location	Gini	S.E.	Lower Bound	Upper Bound
Elementary & Rural	0.002774	0.000229	0.002323	0.003225
Elementary & Urban	0.001991	0.000184	0.001630	0.002352
Gimnazjum & Rural	0.001654	0.000121	0.001416	0.001891
Gimnazjum & Urban	0.002858	0.000279	0.002309	0.003407

Note: The table shows the Gini coefficients of the distribution of leave-out mean of Raven's standardized score at the school level (including individual's class) for combinations of the school type and location.

Since I observe only a sample of classes from a school, the above ranking could be a result of sorting between schools as well as sorting within a school.[17] To disentangle these two effects I run additional regressions described in (2) and (3). Table 4 Panel B presents the estimated coefficients from the regression (2) - β - correlation between a student's and the mean of her classmates' Raven's score. Panel C from the regression (3) - γ - correlation between a student's and the mean of her outside-class schoolmates' Raven's score. The first column shows the results for the whole sample, the second for the urban schools and the last one for the rural ones. Firstly, consider sorting at the entrance to elementary school. Consistently across the sub-samples, the correlation with outside-class peers is smaller than with classmates ($\gamma_1 < \beta_1$), but the difference is insignificant. The hypothesis that $\beta_1 - \gamma_1 = 0$ is not rejected with p-value 7.2% for the urban schools and 9.1% for the rural. This suggests that only sorting between schools is present at the entrance to elementary school.

Before I proceed with the interpretation of changes at the entrance to *gimnazjum*, it is worth to establish a reference point and think what would be the coefficients if *gimnazjum* had no impact on sorting. This happens if there is the same number of *gimnazja* as elementary schools and students from one elementary school are assigned to one *gimnazjum* and they can not request an alternative one. In addition to this, there can not be any adjustment in the class composition (sorting within). If this is the case, changes at the entrance to *gimnazjum* in the correlations with classmates and outside-class peers, will be zero ($\beta_3 = \gamma_3 = 0$). When only sorting within increases: the correlation with classmates increases ($\beta_3 > 0$) and the correlation with outside-

[17]For example, in the absence of change in sorting between, increase in sorting within would on average lead to decrease in correlation with the school level leave-out mean when we randomly draw classes from a school

class peers decreases ($\gamma_3 < 0$); when only sorting between increases: the correlation with classmates doesn't change ($\beta_3 = 0$) and the correlation with outside-class peers increases ($\gamma_3 > 0$); when both types of sorting increase: the correlation with classmates increases ($\beta_3 > 0$) and the change in correlation with outside-class peers is uncertain.

In the whole sample, the change in the correlation with classmates (β_3) is significant but very small in magnitude and the change in the correlation with outside-class peers is significant and negative ($\gamma_3 < 0$), which means that sorting within did not change but sorting between decreases. The effect is potentially large since the correlation with schoolmates in *gimnazjum* decreases by 61%. However the effects are heterogeneous once we look at the sub-samples of urban and rural schools. In the case of the urban sub-sample the correlation with classmates ($\beta_3 > 0$) increases, but there is no change in the correlation with outside-class peers ($\gamma_3 \approx 0$). This shows that sorting within increases and suggests that sorting between increases as well. The change in sorting within is not negligible since the correlation with classmates increases by 23%. For the rural sub-sample, both changes are negative, but the change in the correlation with outside-class peers is more negative than with classmates ($\gamma_3 < \beta_3 < 0$), this pattern can be explained by a decrease in both types of sorting (i.e. students are mixed across schools and classes). The correlation with classmates decreases by 24% and the correlation with outside-class schoolmates by almost 85% (!), which shows a huge drop in sorting between schools. One explanation is necessary here, since there is no evidence for sorting within at the entrance to elementary school, a decrease in this type of sorting at the entrance to *gimnazjum* should be interpreted as a decrease in the homogeneity of students at the class level (students are more mixed within a school).

4.2 The Effect of Competition

The breakdown between the rural and urban areas might be to general to claim about the effect of competition. Therefore in the alternative specification I use a share of non-local first grades to directly relate a change in sorting with a competitiveness of the educational market. In order to control for the potential confounding effect of the school network I include also a ratio of the number of elementary school to the number of *gimnazja*.

Table 6 Panel A presents the results for the leave-out mean at the class level and Panel B for the outside-class mean for different specifications of the regressions defined in (4) and (5). The effect of competition (β_4) is positive and significant in the case of the correlation with classmates (Panel A columns (1) and (3)) but insignificant in the case of the correlation with outside-class peers (γ_4, Panel B columns (1) and (3)). Using similar argument as in the previous section, this suggests that a higher competition is connected with increase in sorting within and sorting between schools.[18] Interestingly, keeping the competition variable constant, an increase in the school network variable has negative but insignificant effect on the correlation with outside-class peers (γ_5, Panel column (3)) and classmates (β_5, Panel A column(3)). Consistently with my previous arguments, the higher the number of elementary school relative to *gimnazja*, the more heterogeneous *gimnazja* are.

Table 6: Regressions Result

	(1)	(2)	(3)
Panel A : 4th Regression: Leave-out mean at the class level			
$\overline{Y}_{-ics}$	.819*** (.02)	.819*** (.02)	.819*** (.02)
$\overline{Y}_{-ics} \times GIM_s$	-.168*** (.044)	.102 (.098)	-.074 (.116)
$\overline{Y}_{-ics} \times GIM_s \times comp_s$	.321*** (0.089)	-	.288** (.089)
$\overline{Y}_{-ics} \times GIM_s \times ratio_s$	-	-.086 (.48)	-.041 (.051)
GIM_s	-.010 (0.009)	-.003 (.009)	-.011 (.009)
n	8882	10496	8882
Panel B : 5th Regression: Outside - class mean			
$\overline{Y}_{-cs}$	.714*** (.054)	.714*** (.054)	.714*** (.054)
$\overline{Y}_{-cs} \times GIM_s$	-.656*** (.141)	.066 (.228)	-.307 (.274)
$\overline{Y}_{-cs} \times GIM_s \times comp_s$	.655 (.347)	-	.51 (.362)
$\overline{Y}_{-cs} \times GIM_s \times ratio_s$	-	-.246* (.105)	-.151 (.11)
GIM_s	-.047* (.022)	-.028 (.021)	-.05* (.023)
n	8067	9612	8067

Note: Robust standard errors in parentheses.*** denotes significance at the 0.1% level, ** at 1% level and * at the 5% level. The outcome variable is a standardized Raven's Progressive Matrix Test score (with subtracted mean and divided by standard deviation).

To summarize, the results presented in Table 4 and 5 show that the inequality ranking between school types, turns upside-down between the urban and rural areas. Moreover, the results suggest that only sorting between schools explains inequalities among elementary schools. The change in sorting at the entrance to *gimnazjum* is heterogeneous.

[18] In the absence of an increase in sorting between, an increase in sorting within a school should lead to a negative correlation with the outside-class peers.

While in the case of the urban schools, sorting between and within increase, in the case of the rural schools both types of sorting decrease. Table 6 shows the alternative evidence that a competitiveness of the educational market is connected with an increase in both: sorting between and within.

4.3 Robustness

One possible concern for the results are the test-room shocks at the time of measurement. Imagine that a barking dog was influencing students' attention during the Raven Progressive Matrix test. Then the correlation might by driven not only by sorting but also by the fact that all students were exposed to the barking dog. Unfortunately, while in elementary schools each student took the test in different times, in *gimnazja* groups of students took the test together. Therefore, the change in the correlations of interests between the stages of education may simply reflect different exposures to the test-room shocks. There are three reasons why this is rather unlikely. Firstly, the measurement was conducted by the team of professional psychometricians with all measures taken to provide neutral environment for all test-takers (Jasinska et al. 2013). Secondly, the nature of these shocks would have to be different between the urban and rural schools, since the changes in the correlations with classmates are different ($\beta_{3,urban} > 0 > \beta_{3,rural}$). I find it rather implausible. Finally, to fully exclude this possibility, I exploit the fact that in almost one-third of *gimnazja* students took the Raven's test in two groups within a class. Thanks to this, I can directly check whether there is any impact on the Raven's score of being in a separate group after controlling for the class fixed effects. The potential significant effect would indicate that the test-room environment matters for the outcome, however the regression shows highly insignificant coefficient, both in the urban and rural areas. On the other hand, the correlation between a 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 (all the results are available upon request).

The other possible explanation are changes in sorting at the entrance to elementary school. Specifically, for the sixth-graders (from 2010) the sorting at their first grade (in 2005) could be different than for the first-graders in 2010. The data limitations does not allow me to fully explore this possibility, nevertheless, the parental questionnaire allows to shed some light on this issue. It asks questions whether a child attended a local, assigned elementary school, which might be considered as a measure of the elementary school selection. Table 7 presents the percentage of parents (of students who are in the first year of elementary school and *gimnazjum*) who answered yes to this question, by the urban/rural breakdown (this question is thus more retrospective for the parents of students from *gimnazjum*). The results show there there is indeed a difference between the elementary school entrants in 2005 and 2010 and students from *gimnazjum* were more likely to go to their assigned school. However the difference is only statistically different from zero in the whole sample (with p-value=4%) and the magnitudes of the difference is very small: 2.9% point for the whole sample, 1.2% points in the case of the rural schools and 4.8% points in the case of the urban. Even though this effect could possibly bias downward the change in sorting between schools at the entrance to elementary schools and *gimnazja*, its magnitude and significance cast doubts on the importance of it.

As for sorting within a school, there are no clear reasons why the principals' practice could change between 2005 and 2010. First of all, the results presented in this paper show that actually sorting within is negligible at the entrance to elementary school. Moreover, there was no reform which would provide additional motivation for student grouping or vice-versa. In addition to this, because the change in sorting within is different in the rural and urban areas, the possible confounding effect would have to affect sorting in a heterogeneous way. I find this possibility rather unlikely.

To summarize, even though more data is needed to fully exclude alternative explanations, there are no convincing evidences that the main results are not because of changes in sorting.

Table 7: Percentage of students who attended a local, assigned elementary school

Stage	All	Urban	Rural	n
Elementary school	79.1%	72.4%	82.1%	7066
Gimnazjum	82%	77.2%	83.4%	4844
difference	2.9*	4.8	1.3	
n	10528	3455	7073	

Note: Percentage of answers "yes" for the question asked to parents whether their child attended a local and assigned elementary school. * - denotes significant difference at the 5% level.

5 Discussion

The results show that the type of the educational market is an important determinants of sorting of students within the Polish comprehensive education. Moreover, the competition between schools seems to be an important characteristic. In the following section I discuss in more detail channels of sorting and possible policy implications. The last part of this section is devoted to the discussion about the causal effect of changing structure of the comprehensive education on sorting of students.

5.1 Sorting Between Schools

There are two main forces responsible for the differential change in sorting between schools: competition between schools and the school network. The later is a likely reason for mixing of students at the entrance to *gimnazjum* in the rural areas. The number of schools coming from Herczynski and Sobotka (2013) and presented in Table 1 shows that in the rural areas there are on average 2.3 elementary schools per *gimnazjum*. Because of catchment areas, students from different elementary schools are directly mixed in one school. Interestingly, this is also the case in the urban areas, where the ratio is around 1.5 but this direct mixing effect is overcome by the competition between schools. In line with the literature on school choice and the effects presented in Table 6, this could explain the increase in sorting between schools.[19] Another (but complementary) explanation is that in the rural areas there is a big change in student mobility between the elementary and the secondary stage of education, while in the urban areas this change is much smaller. As suggested by Kertesi and Kezdi (2013), free school choice and mobility should make residential sorting irrelevant. If the rural areas are characterized by strong residential sorting and no mobility, the composition of elementary schools' students will reflect this. Later, at the secondary education stage, students are more mobile (and there are less schools) so that mixing will take place. In the urban areas, because of higher mobility, residential sorting may have much smaller effect at each stages of education.

Another potentially important thing is to see how the educational market affects the nature of sorting between schools. A change in a school homogeneity can be asymmetrical. On the one hand, sorting could lead to an emergence of elite schools with high-achievers and mixed schools for other students, on the other, to an emergence of ghetto schools which concentrates low-achievers. Knowing, which part of the distribution of student achievement is mainly responsible for growing inequalities might be crucial

[19]For similar evidence in Hungary see Kertesi and Kezdi (2013)

for the proper policy targeting. In order to explore this question, I use the Simultaneous Quantile Regressions to see how the distribution of the Raven's score changes with peer quality. Specifically I run:

$$Y_{ics} = c + \alpha \overline{Y}_{-is} + \epsilon_{ics} \tag{6}$$

where Y_{ics} denotes the Raven's score for student i from class c in school s. $\overline{Y}_{-is}$ is the school-level leave-out mean. I run this regression for the 25th, 75th percentiles and a median. If sorting leads to an emergence of elite schools, the distribution of Raven's score should be less dispersed and more negatively skewed with higher peer quality - this happens if the coefficient α on the 25th percentile is larger than on the 75th percentile, and if the coefficient on median is larger than on the 25th and 75th percentiles.

Table 7 presents the coefficients α from the equation (4) - the correlation between a student's Raven's score and the school-level leave-out mean - for the different school type-location subsamples and moments of the distribution. The last three columns show the differences between the effects on the 25th, 75th percentiles and a median. Among elementary schools, the higher peer quality does not change dispersion of the distribution ($q75 = q25$), however it makes it more negatively skewed ($q50 > q75 = q25$). This pattern is more visible among the rural schools. As for *gimnazja*, the distribution becomes less dispersed ($q75 < q25$) and more negatively skewed ($q50 > q75$), however the effect is larger among the urban schools. These results suggest that high performing elementary schools are not more homogenous than low-performing ones (however they capture relatively more students from the top of distribution). This is not the case for *gimnazja*, where the high performing ones are more homogenous and attract relatively more students from the top. Moreover, while for the elementary education the urban elementary schools are relatively similar to each other (in homogeneity), for the secondary education the results suggest emergence of the elite urban *gimnazja*. The change is not that dramatic among the rural schools.

Table 8: The Simultanoues Quantile Regressions - correlation between a individual's Raven's score and the school-level leave-out mean.

		The Quantile Regression			Differences		
School & Location	OLS	q25	q50	q75	q75-q25	q75-q50	q50-q25
Elementary & Rural	.885** (.023)	.903** (0.046)	1.02** (0.027)	.911** (.033)	.009	-.108**	.116**
Elementary & Urban	.822** (.034)	.822** (0.086)	1.03** (0.098)	.979** (.064)	.157	.156	.208*
Gimnazjum & Rural	.663** (.06)	.759** (0.084)	0.728** (0.067)	.567** (.048)	-.192*	-.161**	-.031
Gimnazjum & Urban	.862** (.022)	1.063** (0.07)	0.947** (0.061)	.686** (.05)	-.377**	-.261**	-.116

Note: The table shows the correlation between a individual's Raven's score and the school-level leave-out mean of Raven's score - this is the coefficient α from the regression: $Y_{ics} = c + \alpha\overline{Y}_{-is} + \epsilon_{ics}$. Robust standard errors in parentheses. ** denotes significance at the 1% level and * at the 5% level.

5.2 Sorting Within a School

Sorting within a school increases in the urban areas and schools which face higher competition. This is in line with a model developed by Epple et al. (2002), which shows that the expansion of private education leads to stratification by ability and motivates public schools to introduce sorting within a school. The reason is that the public school tracking retains high-ability students, which in the absence of tracking, would go to a private school. Without tracking, public schools attract only low-ability students. However, in Poland there is a strong political pressure to equalize composition of students across classes and school principals might don't want to openly brake this rule. Anecdotal evidence suggests that specialized classes might serve as a signaling device. For instance, a French language class effectively excludes the low class students from the peer group and knowing this, the high-class parents might want to send their children to this school and class.[20] The interplay between a competition and sorting within could indicate that *de facto* informal tracking takes place within a school (Figlio and Page 2002), which might reinforce the effect of competition on educational inequalities of opportunities.

The *gimnazjum*'s principals answers to the survey might shed some light on the reasons behind the increase in sorting within. They were asked about the procedures used in the assignment of students to classes. The data has a qualitative nature thus it only provides anecdotal evidences, moreover one has to remember that these are self-reported answers (the reliability of this kind of data is discussed in Betts and Shkolnik (2000b)). Generally, the principals underline that they are trying to create classes with homogenous level of skills and for this purpose they use measures of past achievements (mainly the standardized exam scores after elementary school). Nevertheless, at the same time they tend to sort students based on their language knowledge, sport achievement or special requests from parents and students. Therefore, even though officially they do not track students within a school, in practice they create specialized classes, which are *de facto* forms of tracking.

The political correctness might bias the principals' answers. The questions about attitude toward the external examinations and their usage in various school activities may be a less biased proxy of principals' behaviour. Since there are only 150 *gimnazja* in the sample, one has to keep in mind the results are based on a relatively small sample. Table 9[21] Panel A shows that there is a general pattern that principals from the urban

[20] Also the language knowledge is more diverse in the urban areas, so that there is more need for the class composition adjustment

[21] The exact definitions of the questions are provided in the table notes

schools are more likely to trust and use information coming from the external exams, at the same time, however, they believe that the score matters too much in the educational path of a child. These results are consistent with the hypothesis that practice of merit-based student grouping is more widespread in the cities. However, even though the magnitude is relatively large, the differences are mostly insignificant. The alternative explanation is that the principal's characteristics are different between the areas. The second panel of Table 9 shows that they are almost identical when it comes to the work experience and education[22] but the share of females is higher in the urban areas.

Finally, it would be informative to know whether sorting across classes is connected with any adjustment in teaching resources. Table 7, using the sample of *gimnazja* and regression of the class Raven's average on teacher's characteristics T_{icg} (teacher i who teaches class c in *gimnazjum* g) along with the school fixed effects

$$\overline{Y}_{cg} = c + \kappa T_{icg} + \mu_g + \epsilon_{icg} \tag{7}$$

shows that this is not the case. As a result, it suggests that this form of tracking within a school increases inequalities, since peer effects are not offset by any tailoring of teaching resources.

[22]In practice all principals have the same level of education.

Table 9: Regression of Class-Level Mean of Raven's Score on Teacher Characteristics

	All		Urban		Rural	
Dep.Variable: Mean of Raven's score at the class level						
Teaching Experience	.0009 (.009)	-	.014 (.019)	-	-.006 (.011)	-
1st Teacher Rank	-	.316 (.260)	-	.378 (.688)	-	.286 (.255)
2nd Teacher Rank	-	.373 (.340)	-	.918 (.873)	-	.089 (.262)
3rd Teacher Rank	-	.159 (.347)	-	.361 (.916)	-	.063 (.28)
School FE	yes	yes	yes	yes	yes	yes
n	929	927	403	402	526	525

Note: Unit of observation is teacher-class in *gimnazjum*. Robust standard errors in parentheses.** denotes significance at the 1% level and * at the 5% level. The outcome variable is class mean of standardized Raven's Progressive Matrix Test score (with subtracted mean and divided by standard deviation). Teacher ranks are official professional ranks in the Polish education system; the baseline rank is "no rank"; 1st rank is "contract teacher"; 2nd rank is "appointed teacher"; 3rd rank is "certified teacher".

Table 10: *Gimnazjum*'s Principals

Question	Urban	Rural	Difference
Panel A: Principals and the External Examination			
6th grade exam as a good signal	67.2%	55.6%	11.4
Usage of the 6th grade exam	84.8%	77.8%	7
External examination as a good signal	93.5%	83.4%	9.9*
External examination is random	18%	26.3%	8.3
External examination is too influential	62%	47%	15
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%	10
n	46	104	

Note: 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 significant difference at the 5% level.

5.3 The Causal Effect of Changing the Structure of Comprehensive Education

The focus of this paper is on the interplay between characteristics of the educational market and sorting of student, however it would be also informative to estimate the similar effect of a change in the structure of comprehensive education. The 1999 educational reform has replaced the single 8-year long elementary school by two schools: the new 6-years elementary school and 3-years *gimnazjum*. Knowing the effect of this change on sorting might be crucial for policy-makers in countries, which are considering redesigning and expansion of the comprehensive education.

My research design can provide me with the causal effect on sorting of students, under two assumptions: firstly, there are no changes in a composition of classes and schools during elementary school; secondly, there are no differences between these two regimes in student sorting at the entrance to elementary school. Whereas the first assumption is not very strong, the second is harder to defend. Since in the old regime, sorting of students to elementary school is likely to be higher, one can interpret my result as biased toward higher sorting. The logic is that I compare sorting of students at the entrance to secondary school to sorting at the entrance to elementary school. If the later is the same in both regimes, this provides me with the change in sorting after splitting comprehensive education into two parts.

The bias means that in the case of rural areas I estimate the lower bound of the effect and the actual increase in mixing of student was at least as large as the reported change. Contrary to this, in the case of the urban areas I estimate the upper bound, so in reality the change in sorting might be neutral. Nevertheless, the results clearly show that splitting the comprehensive education is potentially a very effective tool in fighting inequalities of opportunities in areas with low competition between schools.

6 Conclusions

Recently Poland has been shown as an example of a success country when it comes to educational achievements. The Economist (2013) in the article *The best and brightest* from August 17th 2013 argues:

> Poland has made some dramatic gains in education in the past decade. Before 2000 half of the country's rural adults had finished only primary school. Yet international rankings now put the country's students well ahead of Americas in science and maths (the strongest predictor of future earnings), even as the country spends far less per pupil. What is Poland doing right? And what is America doing wrong?

Many scholars claim that the source of the success lies in the relatively long comprehensive education (Jakubowski et al. 2010, Budapest Institute 2014). My paper suggests that the picture is much more complex, as *de facto* we observe different educational systems in the different educational markets. Compare to the initial sorting at the first comprehensive school, the secondary comprehensive school in the urban areas is in reality a form of hidden tracking, while in the rural it is effectively mixing students.

Therefore one has to be careful when drawing a strong conclusions from the Polish case. Also, this paper suggests that the conceptual division between the comprehensive and tracking education might not be enough to fully describe the educational systems.

Moreover, I find evidence that competitiveness of the educational market is connected with higher sorting across classes and between schools. This contributes to the general literature on educational competition as it provides an empirical evidence that competition leads to an increase in sorting within a school. This is consistent with a theoretical model proposed by Epple et al. (2002). Therefore the attention of researchers should focus not only on changes in inequalities across schools or school districts, but also within a school. This further underlines the role of a school principal in levelling the playing field.

More research is needed to draw strong conclusions about the effect of a change in the structure of comprehensive education on sorting of students. Nevertheless, at least in the rural context, splitting the comprehensive education seems to be an effective tool for reinforcing its positive effect on educational equalities.

References

Ammermüller, Andreas, "Educational opportunities and the role of institutions," Technical Report, ZEW Discussion Papers 2005.

Argys, Laura M, Daniel I Rees, and Dominic J Brewer, "Detracking America's Schools: Equity at zero cost?," *Journal of Policy Analysis and Management*, 1996, *15* (4), 623–645.

Betts, Julian, "The economics of tracking in education," *Handbook of the Economics of Education*, 2011, *3* (341-381), 4.

Betts, Julian R and Jamie L Shkolnik, "The effects of ability grouping on student achievement and resource allocation in secondary schools," *Economics of Education Review*, 2000, *19* (1), 1–15.

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

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.

Brunello, Giorgio and Daniele Checchi, "Does school tracking affect equality of opportunity? New international evidence," *Economic Policy*, 2007, *22* (52), 781–861.

Budapest Institute, "Early tracking in education in the Visegrad countries. Policy brief.," 2014. forthcoming.

Bukowski, Paweł and Martyna Kobus, "Educational Equality of Opportunity in Poland," *Ekonomista*, 2012, (5), 605–627.

Dolata, Roman, "Analiza roznicowania sie systemu oswiaty w Polsce na poziomie szkol podstawowych i gimnazjow, edycja 2011," Technical Report, Edukacyjna Wartosc Dodana 2011.

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.

Galindo-Rueda, Fernando and Anna Vignoles, "The declining relative importance of ability in predicting educational attainment," *Journal of Human Resources*, 2005, *40* (2), 335–353.

Gamoran, Adam, "Student achievement in public magnet, public comprehensive, and private city high schools," *Educational Evaluation and Policy Analysis*, 1996, *18* (1), 1–18.

Hanushek, Eric A and Ludger Woessmann, "The role of school improvement in economic development," Technical Report, National Bureau of Economic Research 2007.

Herczynski, Jan and Aneta Sobotka, "Siec szkol podstawowych i gimnazjow w Polsce 2007-2012 - *Draft version of the report*," Technical Report, Instytut Badan Edukacyjnych 2013.

Hoffer, Thomas B, "Middle school ability grouping and student achievement in science and mathematics," *Educational Evaluation and Policy Analysis*, 1992, *14* (3), 205–227.

Horn, Daniel, "Age of selection counts: a cross-country analysis of educational institutions," *Educational Research and Evaluation*, 2009, *15* (4), 343–366.

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.

Jakubowski, Maciej, Harry A Patrinos, Emilio Ernesto Porta, and Jerzy Wisniewski, "The impact of the 1999 education reform in Poland," *World Bank Policy Research Working Paper Series, Vol*, 2010.

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.

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.

Kertesi, Gabor and Gabor Kezdi, "School segregation, school choice and educational policies in 100 Hungarian towns," Technical Report, Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences 2013.

Manski, Charles F, "Identification of endogenous social effects: The reflection problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

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

Pekkarinen, Tuomas, Roope Uusitalo, and Sari Kerr, "School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform," *Journal of Public Economics*, 2009, *93* (7), 965–973.

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

Rothstein, Jesse, "Does Competition Among Public Schools Benefit Students and Taxpayers? Comment," *American Economic Review*, 2007, *97* (5), 2026–2037.

Sacerdote, Bruce, "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly Journal of Economics*, 2001, *116* (2), 681–704.

The Economist, "The best and brightest," August 17 2013.

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

Waldinger, Fabian, "Does tracking affect the importance of family background on students test scores?," *Unpublished manuscript, LSE*, 2006.

Zimmer, Ron, "A new twist in the educational tracking debate," *Economics of Education Review*, 2003, *22* (3), 307–315.