Hidden Tracking or Mixing? The Structure of Comprehensive Education and Sorting of Students in Poland.

Paweł Bukowski*

The length of comprehensive education is at the core of the debate about educational inequalities. However, not much attention has been paid to the design and structure of comprehensive education. Splitting it into two separate schools may either introduce additional stage of sorting (which may have a form of "hidden tracking"') or lead to further mixing of students. In this paper I focus on Poland and compare sorting between schools and sorting within a school at the entrance to elementary school and lower secondary school (*gimnazjum*) which are separate stages of the comprehensive education. To compare general levels of inequalities across different types of schools I regress a student's own measure of background characteristics (The Raven's Progressive Matrices test) on the leave-out mean of her schoolmates. Next, to disentangle sorting within a school and sorting between schools, I compare correlations between a student's own measure and the leave-out means of her classmates and the mean of outside-class schoolmates. Firstly, while in the rural areas elementary schools are more homogenous than *gimnazja*, in the urban areas the opposite is true. Secondly, at the transition between the corresponding stages of education, in the rural areas sorting within and between become weaker, while in the case of urban both types of sorting are reinforced. The likely explanations for this pattern are smaller number of *gimnazja* than elementary schools, the competition between *gimnazja* in the urban areas and "hidden" tracking within schools based on i.e. language knowledge.

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

1 Introduction

Determinants of income inequalities became one of the most important research fields for modern economists and policymakers. Even though there is a plethora of factors shaping inequalities, education seems to occupy the center of attention (Heckman 2011). It is the main channel for the transmission of social status across generations and as it significantly determines wages any inequalities encountered during education are reflected in the future distribution of wages (Murnane et al. 1995). Therefore, it is not surprising to observe a growing interest of social scientist in educational inequalities[1] and especially, the impact of the design of education system.

The length of comprehensive and tracking education[2] is at the core of the debate about educational inequalities (Betts 2011). Firstly, because of peer effects, grouping high-performing (low-performing) students can further increase (decrease) their achievements. Therefore, mixing students should be beneficial for the low-achievers, but harmful for the high-achievers (Sacerdote 2001, Carrell et al. 2011). Secondly, because of the lack of sufficient signals about student abilities, the earlier the selection take place, the more parental background matters for the selection. On the other hand, when grouping is connected with redistribution of school resources and tailoring the teaching process, it may be actually beneficial for all types of students (especially low-achievers) (Kremer et al. 2011). In the European context[3], the literature shows that expansion of a 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). 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.

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

[2]In the comprehensive education the admission process is based on catchment areas. Conversely, in the tracking education, students are sorted into schools or classses based on their merit.

[3]In Europe students are usually sorted into different schools, while in the US - into different classes within a school

However, not much attention has been paid to the design and structure of comprehensive education. Splitting it into two separate schools may either introduce additional stage of sorting (both between and within schools - "hidden tracking"') or lead to further mixing of students. Which effect dominates depends on the educational market and social contexts. 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 higher sorting between schools. Moreover, this may be accompanied by increase in sorting within a school, which overall could reinforce effect of competition on inequalities. Finally, introducing standardized and external examinations should generally increase sorting of students. This is because principals obtain a relatively noiseless measure of student abilities and parents gain access to school rankings (Horn 2009, Van de Werfhorst and Mijs 2010, Bol et al. 2013).

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 components of the reform were the expansion of comprehensive education (from 8 to 9 years) along with the change of its structure. The old (the Soviet type) system which had consisted of eight-year comprehensive elementary school and four-year tracking high school was replaced by six-year elementary school, three-year comprehensive secondary school (*gimnazjum*) and three years of tracking high school. In addition to this, standardized examinations were introduced after each stage of education and became the main admission criteria for schools at the succeeding levels of education.

The main aim of this paper is to evaluate what is the impact of splitting the comprehensive education on sorting of students.[4] I focus on Poland and compare sorting between schools and sorting within a school at the entrance to elementary school and lower secondary school (*gimnazjum*). Both are part of the comprehensive and compulsory education, but are completely separate entities. The admission to these schools is based on catchment areas, however parents may request an alternative school. After each stage of education students are examined using either low-stake (elementary schools) or high-stake (*gimnazja* and high schools) exams. School principals may use this information to sort students between classes, leading to "hidden tracking". While

[4]The counterfactual situation is when there is only one selection stage in the comprehensive education (at the entrance to elementary school). Therefore, under the assumption that there are no changes in a composition of classes and schools during elementary school and no changes in sorting at the entrance to elementary school, this should provide the causal impact of the division of comprehensive education on sorting of students.

in the old system, these processes took place only once (at the entrance to eight-year old elementary school), after the reform they may take place again at the entrance to *gimnazjum*. To compare general levels of inequalities across different types of schools I regress a student's own measure of background characteristics on the leave-out mean of her schoolmates. Next, to disentangle sorting within a school and sorting between schools, I compare correlations between a student's own measure and the leave-out means of her classmates and the mean of outside-class schoolmates.

I find evidence for the heterogeneous effect. Firstly, while in the rural areas elementary schools are more homogenous than *gimnazja*, in the urban areas the opposite is true. Secondly, at the transition between the corresponding stages of education, in the rural areas sorting within and between become weaker (students are more heterogeneous at the class level), while in the case of urban both types of sorting are reinforced. The likely explanations for this pattern are: smaller number of *gimnazja* than elementary schools, the competition between *gimnazja* in the urban areas and "hidden" tracking within schools based on i.e. language knowledge. In addition to this, at the entrance to elementary school, students are sorted between schools but not sorted between classes.

These results show that the design of comprehensive education is not neutral for sorting and thus for inequalities. 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.

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

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 (Tiebout 1956) and send their kids to the same local school. The separation of rich enclaves from ghettos of poor people in modern cities and resulting differences in school composition can be an example of these processes.

The second reason is the school choice. In the Polish comprehensive education system students are allocated to a local school which is obliged to accept any student from a catchment area, however parents may request an alternative one.[5] Lower financial and informational constraints may lead high-class parents to send their kids to a 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. In the case of rural areas, there are not many schools to choose and the potential cost of sending a child to a non-local school is relatively high[6], therefore the educational market is limited. This is not the case in urban areas, where transportation costs are low, number of schools higher and parents more educated.

As a result, these two sources of sorting between schools should cause a positive correlation of background characteristics between students from the same **school**. How *gimnazja* changes intensity of sorting between schools is ambiguous. Firstly, there are less *gimnazja* than elementary schools[7] and since both stages are obligatory, students from different elementary schools will be mixed together in one *gimnazjum*. This decreases sorting between schools. On the other hand, since elementary schools are separated from *gimnazja*, 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.

[5]The alternative school can reject the request.

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

[7]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 *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 kids for whom distance to a school matter more than for older kids.

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). She might want to group students with different characteristics separately into different classes. The most likely reason is willingness to place students from the same area into one class. This practice seems to be advantageous both to parents and principals. Parents can share the cost of transportation[8] or help each other in the case of some school problems (like conflict with a teacher) and principals can coordinate a course schedule with a school transportation, which allows to manage costs more effectively. Nevertheless, because of residential sorting, this will result in grouping similar students together. 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 the 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.[9]

Therefore, sorting within a school leads to a positive correlation of background characteristics between students from the same **class** (but negative with the other classes). *Gimnazjum* should generally lead to 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 (the students' abilities are generally unknown at the entrance to elementary school). Thus splitting the comprehensive education and introducing standardized examination have made much easier for principals in *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, the reform's main aim was to decrease educational inequalities, therefore principals may be facing pressure to randomize class composition.[10]

[8]Then students from the same street have the same class schedule.

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

[10]In the 5th Section I discuss a qualitative study of principals' policies toward a class composition.

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.[11] 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. Beside this, the set of student, parental and school 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 1 summarizes the available sample.

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.[12]

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.

[11] 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

[12] 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).

Table 1: *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.

3.2 Empirical Strategy

In order to see how an additional school within the comprehensive education is associated with changes in general inequalities I run the regression of a student's own Raven's score and the leave-out mean of schoolmates.

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

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 dependent variables: the leave-out mean of Raven's score of classmates $\overline{Y}_{-ics}$ (2) and the mean of Raven's score of outside-class schoolmates $\overline{Y}_{-cs}$(3).

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

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

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 schoolmates (outside one's class) 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 schoolmates. The same logic applies to changes caused by *gimnazja*, Table 2 shows the expected sign of the coefficients β_3 and γ_3 for the combination of changes in sorting.

Table 2: 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	+	-	?

4 Results

4.1 Regression Results

The regressions results, which are presented in Table 3 should be interpreted in a more qualitative way. The sign of the coefficients reflects the direction of changes in sorting, but the magnitude itself is hard to interpret.as there is no natural scale. 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 correlations.

Table 3 Panel A presents results from the first regression (1). In the rural areas[13] 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 2 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 peers quality is the most equal among rural *gimnazja* and urban elementary schools, which are followed by rural elementary schools and urban *gimnazja*.

[13] 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 3: 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)	.59 (.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 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).

Table 4: 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.[14] To disentangle these two effects I run additional regressions described in (2) and (3). Table 3 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 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). Then 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-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 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 there is an increase in the correlation with classmates ($\beta_3 > 0$), but no change in the correlation with outside-class peers ($\gamma_3 \approx 0$) which clearly shows that

[14] 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

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 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 significant sorting within at the entrance to elementary school, 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).

To summarize, the results presented in Table 3 and 4 show that inequality ranking between school types flips 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. While in the case of urban schools, sorting between and within increase, in the case of rural schools both types of sorting decrease.

4.2 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 school 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 the all test-takers (Jasinska et al. 2013). Secondly, the nature of these shocks would have to be different between 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 5 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 change 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 had been 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 5: 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 significant change in sorting at the entrance to *gimnazja* is not heterogeneous. The ranking of school types based on inequalities flips between the urban and rural ares. Moreover, among the urban schools *gimnazjum* reinforces sorting into classes and sorting between schools, while in the rural areas it reduces both types of sorting. What can possibly explain this pattern? In this section I argue that the educational market competitiveness is the main force.

5.1 Sorting Between Schools

Firstly, consider the change in sorting between schools. The effect on sorting between schools in the rural areas can be explained by the general tendency to have a smaller number of *gimnazja* than elementary schools (which lead to student mixing). The number of schools coming from Herczynski and Sobotka (2013) and presented in Table 6 shows that in the rural areas there are on average 2.3 elementary schools per *gimnazjum*. This is the case in the urban areas as well, but generally the educational market is more competitive as there is relatively more *gimnazja* per elementary school (the ratio from Table 6 drops to 1.49), transportation costs are lower and parents more educated. In line with the literature on school choice, this could explain the increase in sorting between schools.[15] Moreover, these arguments are consistent with Figure 1, which focuses on *gimnazja* and shows the relation between share of students from outside catchment area (measure of competitiveness) and the standard deviation of the Raven's score (measure

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

of school's homogeneity). We can see a negative relation between the competitiveness and homogeneity of schools among urban *gimnazja*, and no relation among rural.[16] Another (but complementary) explanation is that in the rural areas there is big change in student mobility between the elementary education stage and the secondary, 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 the 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, the residential sorting may have much smaller effect at each stage of education.

Table 6: 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*.

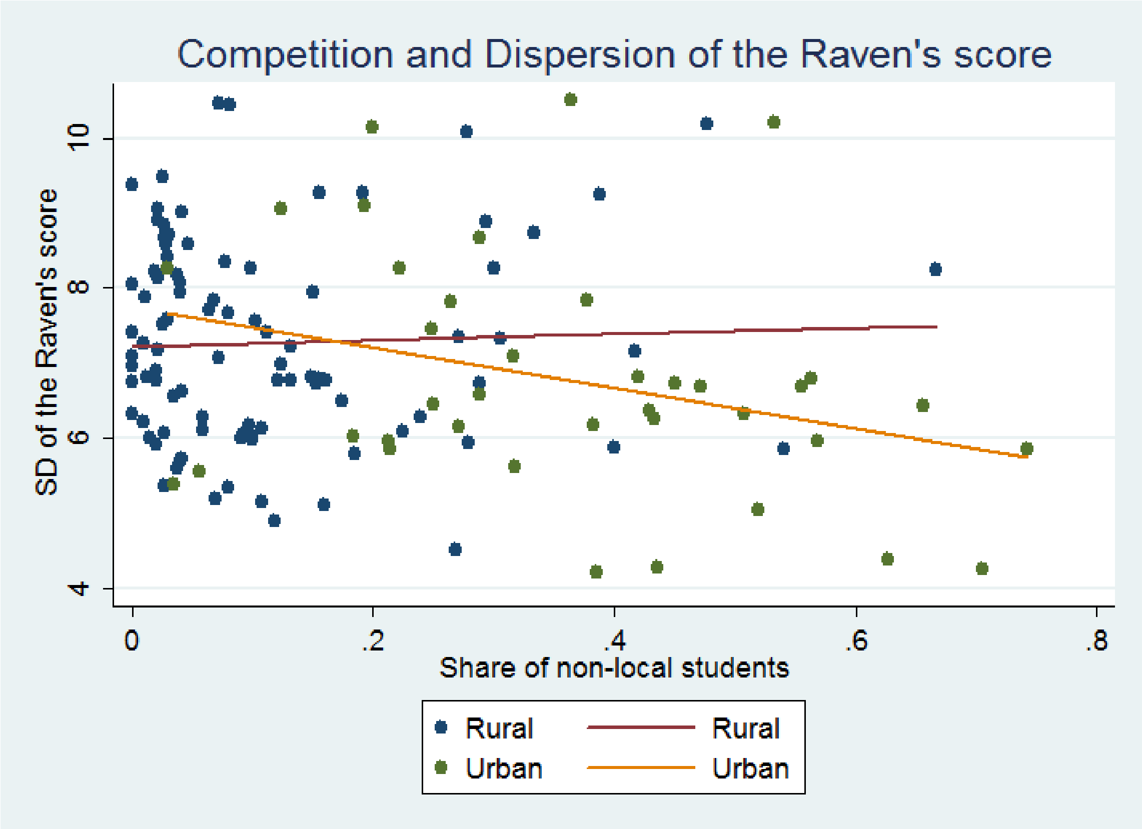
It is important to see how competition affects the nature of sorting between schools. The change in a school homogeneity could be asymmetrical. On the one hand, sorting could lead to an emergence of elite schools with high-achievers and mixed schools for other students, on the other, to an emergence of ghetto schools concentrating low-achievers. Knowing, which part of a distribution is mainly responsible for growing inequalities might be crucial 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:

$$Y_{ics} = c + \alpha \overline{Y}_{-is} + \epsilon_{ics} \quad (4)$$

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

[16] The likely reason why some rural schools have high share of non-local students is that they are located close to metropolitan areas. People who moved out from cities often did not register themselves in the new (rural) place, so their children are considered as non-local.

Figure 1: The Educational Market and Homogeneity of *Gimnazja*.



The figure shows for each *gimnazja* the share of non-local students (X-axis) and the standard deviation of the standardized Raven's score.

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 25t percentile is larger than this on 75h percentile, and if the coefficient on median is larger than on 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 25th, 75th percentiles and 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 7: 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

In the case of the effect on sorting within a school, the likely reason is informal student grouping based on language knowledge or sport abilities. This practice could be more widespread in urban schools, since principals face higher competition and are probably more likely to create specialized classes.[17] If this is the case, it would show the positive interplay between a competition and hidden tracking, which might reinforce the effect of competition on educational inequalities.[18] 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.

One part of the survey asked the *gimnazjum* principals about the procedures used in the assignment of students to classes. The data has 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. As suggested by Figlio and Page (2002), this could indicate that *de facto* informal tracking takes place within a school. Nevertheless, Table 7, using sample of *gimnazja* and regression of a class Raven's average on teacher's characteristics T_{icg} (teacher i who teaches class c in *gimnazjum* g) and the school fixed effects

$$\overline{Y}_{cg} = c + \kappa T_{icg} + \mu_g + \epsilon_{icg} \tag{5}$$

shows that this form of informal tracking is not connected with the adjustment in teachers resources. This may suggest that this form of tracking within a school increases inequalities, since peer effects are not offset by any adjustment in teaching practices.

Hidden tracking might be also reflected in the principals' answers to the questions about attitude toward the external examinations and their usage in various school ac-

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

[18] This will be the case if hidden tracking increases inequalities.

tivities. Since there are only 150 *gimnazja* in the sample, one has to keep in mind the results are based on relatively small sample. Table 9[19] Panel A shows that there is a general pattern that principals from the urban 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[20] but the share of females is higher in the urban areas.

[19]The exact definitions of the questions are provided in the table notes
[20]Practically all principals have the same level of education

Table 8: 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 9: *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.

6 Conclusions

The length of comprehensive education, relative to tracking, seems to be one of the crucial determinants of the educational inequalities. In this paper I argue that its structure is also important and the effect on inequalities depends on competitiveness of the educational market. Using data on Polish students' Raven Progressive Matrix test score I analyze how sorting of students changes at the transition from one stage of comprehensive school (elementary schools) to another (*gimnazja*). I find evidence for the heterogeneous effect. Firstly, while in the rural areas elementary schools are more homogenous than *gimnazja*, in the urban areas the opposite is true. Secondly, at the transition between the corresponding stages of education, in the rural areas sorting within and between become weaker (students are more heterogeneous at the class level), while in the case of urban both types of sorting are reinforced. The likely explanations for this pattern are: smaller number of *gimnazja* than elementary schools, the competition between *gimnazja* in the urban areas and "hidden" tracking within schools based on i.e. language knowledge. In addition to this, at the entrance to elementary school, students are sorted between schools but not sorted between classes. The Simultaneous Quantile Regressions show that the urban elementary schools are relatively similar to each other (in homogeneity), but the high-performing *gimnazja* are more homogenous than the low-performing ones, which suggests emergence of the elite urban .

Generally, at least in the rural context, splitting the comprehensive education seems to be an effective tool for reinforcing its positive effect on educational equalities. Even though the standardized examinations should facilitate sorting, we actually observe that students are mixed, which additionally underlines the importance of the results. This is not the case in areas with a competitive educational market, where both types of sorting increase during the transition to lower secondary school.

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.

Bol, Thijs, Jacqueline Witschge, Herman Van de Werfhorst, and Jaap

Dronkers, “Curricula tracking and central examinations: counterbalancing the Impact of social background on student achievement in 36 countries.,” 2013.

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.

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

Carrell, Scott E, Bruce I Sacerdote, and James E West, “From natural variation to optimal policy? The Lucas critique meets peer effects,” Technical Report, National Bureau of Economic Research 2011.

de Werfhorst, Herman G Van and Jonathan JB Mijs, “Achievement inequality and the institutional structure of educational systems: A comparative perspective,” *Annual Review of Sociology*, 2010, *36*, 407–428.

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.

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.

Heckman, James J, "The economics of inequality," *American Educator*, 2011, p. 31.

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.

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.

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.

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

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.

Murnane, Richard J, John B Willett, and Frank Levy, "The growing importance of cognitive skills in wage determination," Technical Report, National Bureau of Economic Research 1995.

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.

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.