SSE Riga Student Research Papers
2011:3 (132)

SCHOOL-SPECIFIC FACTORS AND EDUCATIONAL OUTCOMES: EDUCATION PRODUCTION FUNCTION APPROACH

Authors: Imants Auziņš
Ilga Bebere

ISSN 1691-4643
ISBN 978-9984-842-43-1

November 2011
Riga

SCHOOL-SPECIFIC FACTORS AND EDUCATIONAL OUTCOMES: EDUCATION PRODUCTION FUNCTION APPROACH

Imants Auziņš

and

Ilga Bebere

Supervisor: Indra Dedze

May 2011
Riga

SSE Riga Student Research Papers 2011:3 (132)

School-specific Factors and Education Outcomes: Education Production Function Approach

Imants Auziņš
Ilga Bebere

Abstract

This study researches the relation between the school-specific factors and the educational outcomes of secondary schools in Latvia. The education production function, which saw its first prominent use in the Coleman Report, is the theoretical basis for this paper. For the estimation of the function at the school level a unique dataset is created. To the best of our knowledge this paper is the first to provide estimates for the education production function at the school level in Latvia. Standardized exam scores are used as a measure of educational outcome. Level of urbanization of a school's location, type of the school as well as the language of instruction shows significant differences in the educational outcomes. Explanations for this are found within the positive and significant pupil-teacher ratio and in the proxy for ability. A discussion of the underlying mechanics is provided and support for the policy of increasing the pupil-teacher ratio as a cost-effective measure of improving educational outcomes are provided.

Keywords: education production function; secondary education; pupil-teacher ratio

Acknowledgements

We would like to express our gratitude to all the people who helped and contributed to our research effort. A special thanks goes to our supervisor Indra Dedze for valuable comments and suggestions for better development of the Bachelor Thesis.

The research itself would not be possible without the data provided by the representative of the Ministry of Education Aija Rudmane.

We also appreciate the support of Uldis Lagzda in obtaining the information for the dataset and the professional help of David Vardanjan in data transformation for the dataset.

We highly appreciate the input of the interviewees to the discussion part of our thesis. Rūdolfs Kalvāns from *State Curriculum Centre* gave valuable insights regarding the general picture of the education system in Latvia. The professional experience of Valentīna Ņefjodova helped us to understand the possible cultural differences between Latvian and minority school students. We are also grateful for the input of Mārtiņš Bērziņš from *Iespējamā misija*, as he was able to provide a different view on the state of basic and secondary education in Latvia.

Finally, we would like to thank the participants of the writing-up seminars at SSE Riga for their valuable comments.

Table of Contents

List of Tables

1 Introduction

Human capital formation has been widely recognized in economics as an important factor for economic growth. Also there is a consensus on the fact that education plays a dominant role in the process of this formation. The contribution of schooling which ends with secondary education to the human capital is more than 50% in Latvia (Ederer, Schuler & Willms, 2007). Therefore, secondary education has an essential role in the economy of the country. Given such contribution the quality of the outcome of secondary education is important in quantitative terms i.e. how good is the outcome and what influences it?

The country of interest for this research is Latvia and specifically the final point of the fully state financed education process – the secondary education. Both basic and secondary education is guaranteed in Latvia by its constitution (The Constitution of the Republic of Latvia, 2010). In the light of this obligation of the state, public expenditure for education is indeed an important policy concern. Still, as education forms only a part of public spending, the question of how much to finance it is ever important. Furthermore, the effectiveness of the financing i.e. where should the funds be allocated to achieve the best result is a significant question.

If one can choose to agree, that it is in the interests of society to have better educational outcomes, then we find it natural to ask what differentiates the educational outcomes? As well we find natural the argument in the public debate, that there should be more financing i.e. more resources for education. Therefore, the question of efficiency of the whole educational system is essential. Still, the literature on economics of education presents globally very varied and even contradicting evidence on these questions. At our present grasp of the literature, we have found no empirical studies conducted specifically in Latvia at the school level, which tackle this issue.

Therefore, in this paper we ask and attempt to answer the question: „What effect do school-specific factors have on educational outcomes of secondary education in Latvia?”. We find this question relevant due to the importance of secondary education in human capital formation and due to the fact that there are only a few studies, which look at the relation between school specific factors and educational outcomes in Latvia. The educational outcomes are measured by indices of standardized exam scores. The approach suggested further allows the analysis of the effects on educational outcomes for both the resources and the school characteristics. For this we employ the method of the education production function first brought forward in the report by Coleman (1966). This approach builds on the analogy of the standard production function encountered in microeconomics and although it is an analogy, it is a well established model in the literature on economics of education. Therefore, we see our work to be predominantly of explorative nature.

First, we review the literature in the field presenting both the theoretical aspects of the model and its development, as well as selected empirical results for it. Secondly, we continue with a description of public education in Latvia. Thirdly, we present our chosen research design for answering the research question. Fourthly, empirical findings are presented, as well as a discussion on these findings, supplementing with qualitative findings from interviews. Finally, we conclude our thesis by answering the research question and reflecting on our work.

2 Literature Review

The aim of this section is to provide an overview of the concepts stated in our research question and our chosen method of linking them together - the education production function and to offer relevant empirical facts from the literature globally and on Latvia.

If we think of human capital as being a resource in an economy's aggregate production function, the productivity of it is important for the total output. Secondary education accounts for more than half of Latvia's human capital (Ederer et al, 2007). Therefore, the quality of this resource or more economically speaking – the marginal productivity of this resource – characterizes the potential total output gains. Hanushek (2006) points out that measurements of cognitive abilities of students do provide important information on the quality aspect. Because of this we ask what determines this quality by looking at the school specific factors. The literature names among others teacher age, pupil-teacher ratio and socioeconomic factors as candidates for determining the educational outcome (Hanushek, 2006; Ammermueller, Heijke, & Woessmann, 2003). Still, to best of our knowledge only Ammermueller et al. (2003) have provided estimates by using the education production function for these factors in Latvia. For this they use the TIMSS data set which sadly has the shortcoming of having many missing values which had to be imputed in the dataset. Therefore, we see it as necessary to make our input in exploring the link between educational outcomes and school specific factors by using alternative sources of data.

2.1 Individual Education Production Function

As stated previously, the education production function has seen its first prominent use in Coleman (1966). His work which is more widely known as the Coleman Report shifted the perspective from solely looking at the inputs of the education system to focusing the attention on the outputs – the educational outcomes, which in turn are explained by the inputs (Hanushek, 2007).

At the lowest level of aggregation – the individual student outcome level – the education production function can be defined in an econometric form in the following way (Hanushek, 2006):

$$O_{it} = f\left(F_i^{(t)}, P_i^{(t)}, S_i^{(t)}, A_i\right) + v_{it}$$

Where “O_{it} – performance of student *i* at time point *t*, $F_i^{(t)}$ – family inputs cumulative to time *t*, $P_i^{(t)}$ – peer inputs cumulative to time *t*, $S_i^{(t)}$– school inputs cumulative to time *t*, A_i – innate ability and a stochastic term v_{it}” (Hanushek, 2006).

Hanushek (2006) also points out two important considerations of this model. First, is the obvious that the student outcome is not only determined by the school inputs, but that the “environmental” factors and the innate ability also have an important role in the educational outcome. The second consideration is of the cumulative nature of the variables. According to him, these are key points to be taken into account when deriving the specifications and also when interpreting them.

As the cumulative nature of the variables may pose considerable difficulties in obtaining the necessary data, Hanushek (2006) also presents the value-added formulation of the education production, which is given below:

$$O_{it} - O_{it^*} = f^*\left(F_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)}\right) + v_{it} - v_{it^*}$$

In this specification the output changes over the period $(t - t^*)$ are related to the changes in the inputs (note, that the ability term cancels out in the individual student case, if it is assumed to be constant for the student). The resulting flow model lessens the data requirements as also variables, which are not in cumulative form, can be employed.

2.2 Aggregated Education Production Function

What essentially makes the aggregated form of the education production function different from the micro-level function is the usage of some kind of average educational outcome (e.g. mean of the standardized test scores as used for example in Andrews, Fayissa & Tate (1991)). This also permits the usage of more easily accessible aggregated measures to proxy peer, family and ability effects, which due to obvious considerations are mostly costly and hard to obtain. Still, the aggregation approach has its own shortcomings which have to be considered.

Hanushek, Rivkin & Taylor (1996) raises the debate for two issues which are present in aggregation. The first stems from the ever possible model misspecification, and the second speaks about the possible benefits of reducing errors-in-variables through higher levels of aggregation.

According to him the misspecification problem becomes dire in the case, where the aggregation level specific effects (e.g. state or school specific effects) are not taken into account (Hanushek et al., 1996). An example would be that a family would choose to live in a district with better educational establishments, if it cares more for the education of its offspring. Therefore, not accounting for the choice of settlement may bias the coefficients on the variables of interest. Thus, to avoid this, specific regional effects should be accounted for.

Also they do not find confirming evidence for the logic that aggregation reduces errors-in-variables. In his work it is restated that the lower the level of aggregation for the explanatory variables, the more reliable are the estimates procured (Hanushek et al., 1996).

2.3 Summary of Findings in the Literature

Except for the Coleman Report (Coleman, 1966), most of the studies in this field have been opportunistic in their approach i.e. using the data (mostly administrative) which has not been compiled for specific research purposes (Hanushek, 2006). In the same source Hanushek (2006) summarizes the coefficient signs and significance for 376 studies performed in the U.S. Without reporting the whole summary of those studies Hanushek (2006) directs the attention to the important point that much of the research reports insignificant coefficients for real classroom resources (e.g. teacher-pupil ratio), financial aggregates (e.g. expenditure per student) and other factors. Although theory predicts a positive sign for all of these measures, the evidence is not that convincing.

Nonetheless the aforementioned review of Hanushek (2006) shows that in the case that the estimates are statistically significant, they are mostly positive as predicted by theory. Therefore, it is of interest to see what results can be procured in the setting of Latvian schools.

Apart from Hanushek (2006) which reviews mostly studies from the United States, we also think it is interesting to look at the evidence in other parts of the world. The East Asian region has been looked at by Woessmann (2003a). He finds that smaller classes (lower pupil-teacher ratios) have a statistically insignificant result on performance, as well as shows that the effect becomes more dubious when controlling for family background. The level of urbanization of the school’s location is important for some of the countries surveyed, whereas it is contrasted with Japan, France and Spain for which these factors carry no significance.

Also, the education of teachers seems to carry no importance. Although on average larger classes in East Asia are associated with higher performance, the author finds an explanation for this phenomenon in the fact that in Japan, Singapore, Hong Kong and also in the comparative case of France low-performers are put in smaller classes. Such educational practices become of importance when evaluating the impact of school-specific factors.

One interesting aspect of the teachers' characteristics is mentioned in Woessmann (2003c), where the author finds a significant negative relationship between teachers' age and student performance. As this finding contradicts the common sense belief that experience comes with age, we find it necessary to include this into our model.

The closest empirical evidence, which can be found for Latvia, comes from Ammermueller et al. (2003). This study estimates the education production function at the individual student level for seven Eastern European countries including Latvia. The educational outcome date comes from the Third International Mathematics and Science Study (TIMSS); therefore, test scores for math and sciences are used as outputs. Main variable of interest for this study is class-size from the category of school resources. The evidence for Latvia presents no significant results for school factors, except for a negative effect of class-size for science scores. Most of the individual characteristics are significant. Family factors, namely attainment of tertiary education for both parents, are of significance, as it is generally true in the U.S. settings (Coleman, 1966). Still, the findings in this paper have to be treated with a healthy deal of care, as the authors admit that half of the data for Latvia was imputed as it was missing.

An approach similar to the education production function has been explored by Geske, Grinfelds, Dedze & Zhang (2006), where individual educational outcomes are explored by the degree of urbanity of the attended school's location, the socioeconomic status and individual characteristics of the students. They find high regional disparities between Riga and other urban areas with respect to rural areas. Also, the distribution of grades varies significantly according to the regional variable. They find an explanation for these differences in the family backgrounds of the students. An intriguing point for the family background comes from Henderson & Mapp (2002). They find that parent involvement improves student achievement and that it varies among ethnic groups, which is a relevant consideration for the case of Latvia we are examining.

An interesting consideration for the urbanization finding of Geske et al. (2006) is the observation of Hanushek et al. (1996). Hanushek et al. (1996) find that pupil-teacher ratios are more likely to become "irrelevant" (insignificant) in lower levels of aggregation of student performance. The authors explain this by within-state and within-district differences as the spending and financing structures are likely to be different. Although this is an interesting point and it appears to be relevant for the U.S. case, we show further in this paper that the education system in Latvia does not allow for great differences in the different regions.

Woessmann (2003b) also argues that it is necessary to control for ability of the students in the case, where it cannot be assumed that it is the same on average. This gives further implication for our model to account for, as it is unreasonable to assume that schools do not look for more talented students for their programs. Anecdotic evidence for this is provided by the different entry requirements for the school programs. There are schools which have a tough competition for students to get in.

3 Public Education System in Latvia

According to the Education Law (2011), four levels of education can be distinguished, namely pre-school education, basic, secondary education and higher education. The first three stages are stated to be free of charge and every person in Latvia is entitled to obtain education up to and including secondary level. Higher education, on the contrary, often includes tuition fees, but as it is out of the scope of this paper, a more detailed discussion on the higher education here will not be provided. Also, we focus only on publicly provided education disregarding the private schools in Latvia.

Pre-school education should be undertaken before or at the age of 5 when it is mandatory. Basic education (grades 1 – 9) is also mandatory. Yet, in the case the person does not obtain basic education before turning 18, the attendance becomes discretionary. The next step includes 3 or 4 years of studies at a secondary or a vocational school. This level of education is voluntary.

At the same time exams for obtaining the diploma for secondary education are the only standardized exams in Latvia that makes these examination results comparable across the country and relevant for our research.

The content of these exams is in Latvian and that raises the question about the importance of the language of instruction. Language of instruction is one of the key characteristics of the education system in Latvia, which has been widely discussed in the public (I. Dedze, personal communication, December 29, 2010; Kuzmina, 2010).

The role of the Russian language of instruction in schools in Latvia has always been a hot topic. In 2004 a reform was done and since then at least 60% of all the subjects in minority schools should be taught in Latvian and the other 40% are taught in the language of minority at the level of secondary education. As most of the minority schools in Latvia use Russian as the language of instruction (see Appendix 1 – Sample Statistics), Russian is used as the language of instruction for those 40% of subjects. (Amendment of Education Law, 2004). The results of this reform are unequivocal, as the representatives from Ministry of Education and Science say that the results of standardized exam for natural science subjects have improved for minority students (Kuzmina, 2010) during the last couple of years. Yet, it might just show that students have got used to the fact that all the exam problems are in Latvian, as it is the case since 2007. Similarly, the ministry admits that the actual execution of the amendments in law is poorly controlled (Kuzmina, 2010). Recently one of the political parties suggested even more strict regulations regarding the language of instruction in schools, as they prefer state-financed school moving towards teaching only in state language (Kuzmina, 2010), but until now no legislative changes have occurred. Also the argument here seems to be more of a political nature.

The financing of schools in Latvia is arranged in the way that part of the funding comes from the state and the rest is financed by the resources available in the respective municipalities. The government transfers the funds for teacher salaries according to the regulations set by the Cabinet of Ministers of the Republic of Latvia. The funding assigned from the state budget varies from year to year as it is based on the resource allocation chosen for the particular year's budget. For instance, when developing the budget for 2009 the funding for schools was reduced by approximately 50% (The Ministry of Education and Science, 2009). Finally, the director of the school decides how to allocate the money within the educational institution. This person then can affect the study outcomes e.g. paying out bonuses for teachers and using other types of motivation.

Mainly due to the current economic crisis and the problems it highlighted in the structure of public sector in Latvia, the education system recently underwent certain transformations. According to the World Bank and the International Monetary Fund, there are possibilities to increase the efficiency of the education systems by increasing pupil-teacher ratio and introducing a new mechanism for calculating salaries for teachers, which was introduced in 2008 (The Ministry of Education and Science, 2009). The new mechanism is based on the voucher system that takes into account the number of pupils in the school rather than the number of teachers, as it was done before. Other components of the new system for the purpose of calculation of the teacher wages take into account the grade level the teachers are serving and the type of educational institution they are working in. For instance, teachers in state gymnasiums earn 10% more than those working in ordinary secondary schools (Cabinet Regulations No. 1616, 2009). All the calculations take into consideration the estimates that there should be approximately eight pupils per teacher in rural areas and about 10 pupils per teacher in urban areas (Cabinet Regulations No. 1616, 2009). Latvia currently has one of the lowest pupil-teacher ratios among the EU member states (Eurostat, 2010); therefore, the government aims to increase this ratio even further as that allows reducing costs (The Ministry of Education and Science, 2009).

There are some struggles regarding the efficiency of the voucher system, because some rural schools can be in danger as the small number of students they have yields less funding (Dreijere & Valaine, 2009). The alternative is closing these educational institutions, but that would lead to limited access to the schooling in rural areas that is already characterized as being, on average, of lower quality (Geske, Grinfelds & Kangro, 2003). Thus, the government should find the balance between the expenses and education efficiency in the form of study outcomes and education availability.

4 Methodology

The analysis will be based on the education product function widely discussed in the literature, but the regression equation will be modified according to the unique characteristics of the education system of Latvia.

The aim of this study is to identify the impact of changes in various education performance determinants on the study outcomes, and looking at school level seems to be the most appropriate choice. It allows analyzing the overall situation by looking at the individual school-specific factors. We want to highlight the impact on education outcomes at school level as it should help to define what actual problem the education system faces in relation to its availability and equality. Also we opt to exclude demographic and behavioural patterns that are hard to observe; rather the error term contains these features.

4.1 The Method

The education production function is the theoretical framework for this study. Thus, the primary method of answering the question for the school specific factors is employing econometric estimation of the said function. As secondary education in Latvia is predominantly publicly provided, there exists a clear system of provision. Therefore, we find it reasonable to look at the parameters of the system and to see whether the variation in the school specific factors can explain the student achievement differences at the school level.

In the literature review we have mentioned, that such estimates are not very consistent across different studies. To our present knowledge, no education production functions have been done previously for public secondary education in Latvia. Therefore, it poses a challenge to our potential results in terms of reliability. To somewhat counter this, we plan to conduct interviews with the stakeholders of the educational process. In this way we enable ourselves to combine both the insights of the quantitative analysis - the statistical significance of the different factors together with expert knowledge. This can be seen as methodological triangulation as defined in Mathison (1988). This enables to either support the quantitative findings, thus resulting in convergence of them or indicating contradiction i.e. divergence, therefore uncovering possible flaws in the primary approach and maybe suggesting additional improvements to the model.

4.2 The Model

With this section we would like to bridge from the general model described previously (Hanushek, 2006a) to the model we employ in our work. First issue to be addressed is to move up in the level of aggregation from the individual pupil educational outcome to the school level educational outcome. The model we plan to use can be represented in the following form:

$$O_{it} - O_{it^*} = \alpha + \beta_1 S_i^{(t-t^*)} + \beta_2 A_i^{(t-t^*)} + \beta_3 D_i + v_{it} - v_{it^*}$$

Where the *O* difference is the difference in the average standardized exam score for the school, alpha (α) is the constant term, *S* difference is the vector of differences in school specific factors and *A* difference is the difference in ability. *D* denotes dummies which control for differences such as rural/urban, language of instruction and others which will be discussed in the next sections.

This model deviates from the general one presented previously in the following ways:

1) Family inputs are disregarded as the data limitations do not allow to control for it. Still, we consider them in our qualitative data in the discussion part of the paper.
2) Peer inputs are disregarded. We find it justifiable to do this due to the same logic ability has been disregarded in the general model presented by Hanushek (2006) i.e. the peer characteristics on the school level are regarded to be constant across time, therefore the difference approach cancels the term out.
3) The ability term is reintroduced as a difference to allow controlling for the changes in it.

For the estimation of this model we employ a panel data set and use the fixed-effects model approach. The absorption will be done on time. The fixed-effects enable us to capture time effects, so that we can estimate the coefficient on the variable of interest – namely the school specific factors. The coefficients will be estimated using OLS and the model itself is in linear form. The usage of the linear form has been justified in Boissiere (2004) as a reasonable approximation resulting from the standard Cobb-Douglas production function.

4.3 Data

For our regression purposes we work with the data set that was created by combining several information sources. Data on standardized exam scores, instruction language, location and school type was obtained from the *State Curriculum Centre (SCC)* (*Valsts Izglītības Satura Centrs*). Here some information was available for 6 years, whereas the dataset that can be reproduced for our specifications was presented only for the last two study years (2008/2009 and 2009/2010). This obstacle shaped the further creation of our dataset. The Ministry of Education and Science of the Republic of Latvia released additional information on other items of our dataset, namely the number of teachers, their characteristics and overall number of students studying at every school. The data provided by the ministry was already combined at the school level, whereas we manually had to work on creating the average school-specific values for exam scores. The developed dataset was further supplemented by a rating of student performance in state olympiads, which we use to proxy ability. State olympiads are organized to offer students a chance to prove their knowledge and ability in a competitive setting. This rating has been compiled by the foundation *Draudzīgais aicinājums*. After the data transformation we obtained an unbalanced panel with 233 observations (out of 377 schools in total) for academic year 2008/2009 and 343 observations (out of 375 schools in total) for years 2009/2010.

4.3.1 Standardized Exam Results

There are several state tests and exams employed in basic and secondary grades in Latvia, but as our main concern is education outcomes for secondary schools, the results of standardized exams should serve as a good proxy for study performance of students. Thus, we define the standardized exam results as our first key variable.

The data for standardized exams is available on the webpage of the *State Curriculum Centre*. The *SCC* collects the exam results since the academic year 2004/2005. There are 11 standardized exams: English, biology, physics, French, chemistry, Russian, Latvian for minorities, German, mathematics, history, Latvian. We concentrate on 3 different measures for educational outcomes, namely overall weighted average grade, weighted average grade in natural sciences and the weighted average grade in humanities. The weights are the number of exams taken in the respective subject by the students of the school. The mean school scores are created in the following manner. We base the values on the total percentage obtained in particular exam. For instance, we take all the English exam writers from the same school, count together their percentage points and divide by the number of students taking the exam. Basically, we calculate the mean values for each exam within each school for every observation period.

For analysis of certain specialization subjects, we create two new variables. One includes average grades for mathematics, physics, chemistry and biology and describes the natural science study outcomes, whereas humanities include various language classes and history. It is argued that the number of students interested and successful in natural sciences is low in Latvia (Iespējamā misija, n.d.; Vincevs, 2006), and we want to see whether the influence of the school-specific factors employed differs among the mentioned subject areas.

4.3.2 Urbanization

According to analysis based on PISA and TIMSS studies, strong disparities in educational achievements exist if one compares urban and rural schools (Geske et al., 2003) in Latvia. We expect to find similar patterns when using standardized exam results; therefore, we define a variable that accounts for school location. Based on the data obtained from the *State Curriculum Centre*, we will distinguish four situations. School can be situated in Riga, Republican city, a smaller city or in the countryside. This way we can evaluate whether there is also a statistically significant difference in standardized exam results among the urban schools. The different levels of urbanization are categorized according to the Law on Administrative Territories and Populated Areas (2011) with the exception that Riga is taken as a separate case.

4.3.3 Type of Educational Institution

We exclude vocational education from our regression specification due to the fact that they provide predominantly vocational secondary education in contrast to general secondary education; however, we prefer to check for the differences between secondary schools and gymnasiums. Education obtained in gymnasiums or state gymnasiums is usually associated with better study outcomes in Latvia; therefore, this parameter might be considered to be country-specific. We expect better results for state gymnasium students as they receive the status only after attaining certain study outcomes (certain percentage of excellent grades should be accomplished) (Cabinet Regulations No. 129, 2001), but the results are not unequivocal because we cannot exclude the possibility that certain schools prefer to remain as secondary schools or gymnasiums. Insufficient number of students in upper classes prevents the granting of state gymnasium status even if the study results are high (Cabinet Regulations No. 129, 2001). A variable distinguishing the three school statuses is introduced for this analysis. We include also five boarding and magnet schools. Due to data limitations we look at them jointly. Boarding schools are schools where some or all of students live during their school years together with their fellow students in a dormitory type of setting. Magnet schools in turn are schools which offer specialized curriculum. In the case of Latvia such schools most often accommodate students with different kinds of disabilities.

4.3.4 Language of instruction

Latvian and Russian are the main languages of instruction in Latvia; yet, other languages such as Polish, Estonian, Lithuanian and others are sometimes used. We construct variables from *SCC* data that include the information on this specification. It has four different states, namely schools with Latvian, Russian, mixed or other minority language as the language of instruction.

We expect that in case of significant impact on the standardized exam results, variable might have caught some cultural differences among the different ethnicities. The differences in the perception of education across different ethnicity groups have been discussed in Geske et al. (2003). Nevertheless, the authors elaborate that exam results between schools with Russian and Latvian languages of instruction are rather similar (Geske et al., 2003). It is also worth mentioning that we make the assumption that schools with Latvian as the language of instruction are predominantly attended by students with Latvian ethnic background and schools with Russian as the language of instruction are attended by students with Russian ethnic background or Russian-speaking background.

4.3.5 Pupil-Teacher Ratio

The Ministry of Education and Science has provided information on total number of students studying every year in each school as well as the number of teachers in each school. Using these two values we calculate the pupil-teacher ratio for each school for every observation year. We anticipate that this variable might provide some insights regarding the efficiency of voucher system recently employed in Latvia. Similarly as in the study by Levacic, Jenkin, Vignoles & Allen (2005) it was concluded that reducing pupil-teacher ratio improves study outcomes in England. In Latvia we would expect smaller classes having on average a positive impact on study outcomes *ceteris paribus*. Yet, we should be aware of that in Latvia smaller classes are usually either in private or rural schools. Therefore, we need to look at the interaction of urbanization and pupil-teacher ratio.

4.3.6 Teacher Education

Even though the statistical significance of teachers' education is often questioned in previous studies (Woessmann, 2003b; Buddin & Zamarro, 2009), it can be argued that this characteristic can be seen as one of the aspects that shape the quality, school prestige and students' study outcomes. Higher level of education normally should represent higher competencies in the corresponding field; therefore, we would like to examine whether this factor does have a statistical importance in case of Latvia. In the data we have procured we see eight different entries which describe the education of teachers in the school. The eight variations are in line with the requirements set by the Latvian government (Cabinet Regulations No. 347, 2000). It shows the amount of teachers which have a specific level of educational attainment. Sadly, we do not find any reliable way in the literature or in discussions with experts to transform this data into a variable which can be incorporated into our model. Therefore, we opt to exclude this factor from our specification.

4.3.7 Teacher Age

Similarly, teacher age can be treated as a factor affecting educational outcomes but more from a psychological perspective. Students might feel more open-minded in communication with younger teachers, thus, being more active and interested in lessons, while more senior teachers can obtain the desired outcomes through respect and experience. We would like to test whether the teacher age on average affects the educational outcomes. The raw data includes categories based on five-year age spans, meaning that the first category includes people of age 24 or younger; the next group is for 25-29 years old and so forth. From this information we define a variable that operates as an approximate measure of individual school's average age of teachers. We take the middle value of every time span (27 for age span 25-29 and so forth); to keep the same intervals, we use value 22 for the first category and age of 67 for the last. Weighted average formula is then used to calculate the final numbers that range from 33.94 to 52.93 years.

4.3.8 Results of State Olympiads

Results of state olympiads is a proxy for the aggregate level of ability within a school; thus, we want to check for the relationship between olympiad results and exam grades. The measure with which to approximate the impact of ability is the rating according to which the *Draudzīgais aicinājums* Award is granted. A special ranking developed in 2000 accounts for received places in state olympiads in various subjects, where better results lead to a higher position in the ranking. The ranking is based on olympiads for students starting from the eighth grade (only one subject is represented in eighth grade, namely Latvian language), while the majority of these intellectual competitions are for secondary school students (Draudzīgā aicinājuma fonds, 2009). The ranking is calculated on a yearly basis and the best performing students and corresponding schools and teachers receive medals and other attributes of intellectual testimony (Draudzīgā aicinājuma fonds, n.d.). The *SCC* also collects the statistics on the results of state olympiads, but the time span is not as wide as for the *Draudzīgais aicinājums* Award and the applicability is questionable as no single index of performance is produced. The problem is that the Award was created for a ten year period, meaning that the prizes were given for the last time in 2010 and the choice of winners was based on prior performance rather than the success in the particular academic year (Jankelevica, 2010). Now a new improved model is under construction. This situation reflects the missing values on the second observation period in our regression.

We are aware that the state olympiad results might have an ambiguous effect, because the number of highly intelligent people studying at a particular educational institution might be correlated with the overall performance of a school. Probability that all the students with the highest potential will participate in the olympiads can also be questioned, because there are certain participation limits as well as internal competitions within schools. Still, we opt to include this variable, because it allows us to control for the ability level at an individual school.

4.4 Interview Design

Qualitative method in form of interviews is used to supplement the findings drawn from the estimates of the education production function. The questions are mainly open-ended in order to obtain more in-depth opinions. The main points of interest are current trends, challenges and drawbacks of education in Latvia. Our regression results (in a qualitative way) will also be presented during the interviews to find possible explanations from the expertise of the interviewees. For the list of beforehand prepared questions see Appendix 3; however, additional questions emerged during the actual communication based on the answers and those are not reported. Here we have to note that separate section for interview results will not appear, these findings will rather be incorporated in the discussion part.

Concept of stakeholder analysis is employed when searching for the best respondents for the interviews. We have agreed to interview three people representing different levels in the education system in Latvia. The first interview was conducted with Rūdolfs Kalvāns, director of the division in *State Curriculum Centre* that is responsible for successful realization of all examination at basic and secondary education. He also runs two courses in University of Latvia, in Faculty of Education and Psychology and is in the process of writing his PhD Thesis on education. Together with experience in teaching and performing administrative tasks at school his competence is of considerable merit, thus allowing us to enrich our discussion with a professional opinion.

By employing snowball sampling discussed in Bryman & Bell (2003) we approached a secondary school teacher Valentīna Ņefjodova who currently teaches Russian language in Riga Secondary School No. 64, but she also has an experience in running Latvian classes in schools with Russian as the language of instruction. The acquaintance with both Latvian and Russian students can give us extra insights regarding the possible discrepancies among schools of different languages of instruction.

The last stakeholder of our interest is the organization *Iespējamā misija* (Mission Possible) that offers recent university graduates to involve in teaching process and promote rather practical and hands-on approach to learning (Iespējamā misija, n.d.). Mārtiņš Bērziņš, graduate of this two-year program and current mentor of new participants, agreed to take part in the discussion about the education system in Latvia, its recent trends and problems.

5 Empirical Findings

Before running the regressions we set the Latvian language of instruction and an ordinary secondary school that is located in Riga as the base case. These variables will not be used as independent variables; the intercept point will describe the average effect on a school which is in Riga, is a secondary school; with Latvian serving as the instruction language, holding everything else constant. We choose these particular variables as those are the most frequent cases in the dataset. Of course, alternative choices would not affect the overall findings. This setting will apply to regressions (1) to (3), while for the last two specifications interaction terms are introduced. The regression output for the average results is summarized further under this section in a table. Regressions on the results in natural sciences and humanities can be looked up in the respective tables in Appendix 2. Unless otherwise noted, we refer to the *averagegrade* specification.

Also we should note that all the regressions which included observation from both years were run with absorption on time. The time factor was insignificant with a p-value of about 30%.

The results where the average of all the standardized exam grades is used as the dependant variable should represent the most general view as the impact of every written exam is included. Therefore, we focus more on the output of this regression (see Table 1). We first examine the coefficients on the average teacher age, pupil-teacher ratio and state olympiad results (regression not reported in the table). The values are in line with the findings from other regression specifications that the teacher age on average has a statistically insignificant positive effect on exam grades, keeping everything else constant. This might stem from the reason that the disparities in average age for various schools are not large. The average value is 45.37 years and standard deviation is 2.83. Therefore, we have a rather small subset of possible age values to draw inference from.

The coefficient on pupil-teacher ratio, on the contrary, is a strong predictor of an increase in exam grades even at 1% significance level. Only the magnitude of the effect appears to depend on whether we control for ability in the form of the state olympiad results. As it can be seen in the table the values are lower if we exclude the state olympiad variable (regressions (1), (2) and (4)) and a one point increase in the ratio (one more pupil per teacher) leads to a corresponding 0.466, 0.377 or 0.437 average increase in the final exam score at the school level.

Dependent variable:	***Average exam grade***				
Regressor	(1)	(2)	(3)	(4)	(5)
Average teacher age	0.148 (0.124)	0.189 (0.125)	0.107 (0.191)	0.181 (0.124)	0.208 (0.181)
Pupil - teacher ratio	0.446*** (0.085)	0.377*** (0.082)	0.854*** (0.220)	0.350*** (0.102)	1.483*** (0.398)
Gymnasium	6.414*** (0.942)	5.415*** (1.001)	5.686*** (1.275)	5.497*** (0.993)	5.881*** (1.234)
State gymnasium	12.250*** (0.901)	11.862*** (0.986)	9.158*** (1.786)	11.902*** (1.010)	9.504*** (1.746)
Boarding and magnet school	-5.546** (2.758)	-6.618** (3.084)	-3.649 (4.721)	-6.433** (3.126)	-2.950 (4.773)
Russian language	-1.314 (1.012)	-3.168*** (1.144)	-3.042 (2.049)	-3.116*** (1.138)	-3.284 (2.046)
Mixed language	-1.091 (1.185)	-1.376 (1.342)	0.155 (2.027)	-1.414 (1.353)	0.592 (2.155)
Minority language	10.676 (3.168)	7.577*** (2.947)	13.471*** (3.203)	7.509*** (2.896)	14.971*** (3.683)
State city	-	-2.184* (1.203)	-2.338 (2.028)	-2.960 (2.671)	-0.208 (7.710)
City	-	-3.558*** (1.115)	-3.597* (2.051)	-3.925 (2.818)	5.972 (7.559)
Rural area	-	-4.401*** (1.166)	-3.006 (2.041)	-6.310** (2.518)	7.481 (5.402)
Pupil-teacher ratio × *State city*	-	-	-	0.060 (0.192)	-0.225 (0.672)
Pupil-teacher ratio × *City*	-	-	-	0.030 (0.214)	-0.931 (0.695)
Pupil-teacher ratio × *Rural area*	-	-	-	0.197 (0.204)	-1.067** (0.480)
State Olympiad results	-	-	0.660*** (0.153)	-	0.638*** (0.158)
Intercept	38.699*** (5.828)	40.969*** (5.783)	38.270*** (9.002)	41.664*** (5.768)	27.114*** (8.940)
F-statistics for the hypothesis that the population coefficients on the regressors are all zero:					
School type	72.46 (0.00)	52.75 (0.00)	13.47 (0.00)	50.77 (0.00)	14.96 (0.00)
School location	-	5.13 (0.00)	1.05 (0.37)	2.17 (0.09)	0.86 (0.46)
Language of instruction	4.75 (0.00)	6.07 (0.00)	9.24 (0.00)	6.06 (0.00)	9.72 (0.00)
Pupil-teacher ratio × *School location*	-	-	-	0.31 (0.82)	1.94 (0.12)
Regression Summary Statistics					
R^2	0.278	0.302	0.340	0.303	0.358
$\bar{R}^2$	0.266	0.287	0.304	0.284	0.313
n	576	576	233	576	233

Table 1 – Regression results using *averagegrade* as the dependant variable. Source: made by authors

Note: All regressions include an intercept. Heteroskedasticity-robust standard errors are given in parenthesis under estimated coefficients. *p*-values are given in parentheses under *F*-statistics, which is heteroskedasticity-robust. Coefficients are significant at the *10%, **5%, ***1% significance level.

As we have data for state olympiads only for one observation year, the regression specifications with the ability proxy in form of the state olympiads rating include the data only for one year. Thus, one can observe the decrease in the number of observations in the table. When accounting for student ability, the effect of changes in the pupil-teacher ratio is more pronounced. It might reflect that the class size may play a bigger role for academically more talented students.

Regressions (4) and (5) display differences in effect of extra student per teacher among various urbanization levels. Even though the positive effect of an increase in pupil-teacher ratio appears to be smaller outside Riga in the specification (5), statistically significant discrepancies appear only in rural areas, where increase in standardized exam grades is on average 1.067 points lower than among capital schools, holding everything else constant. If we exclude the state olympiad results, the signs for the interactions of pupil-teacher ratio and urbanization dummies take the opposite value and statistical significance disappears. Similar findings can be drawn when analyzing natural science grades (see Appendix 2). Still, the interactions of pupil-teacher ratio and urbanization remains significant for city and rural area schools when having humanities as the dependant variable.

The state olympiad variable itself carries a statistical significance at 1%. A one point increase in the ranking by *Draudzīgais aicinājums* foundation leads to an average 0.66 point increase in the average performance of a school (regression (3)), holding everything else constant.

In regressions (2) and (3) we compare three groups of dummy variables. The categories are based on school type, instruction language and urbanization. Regressions (4) and (5) are supplemented by group of interaction dummies representing the effects of alteration in pupil-teacher ratio across various locations. The coefficient on state gymnasiums is always strongly positive with the peak impact in regression (1) when the average grade for a state gymnasium compared to a secondary school leads to an average difference in standardized exam scores of 12.25 points. This situation can be explained by the reason that the variable for state gymnasiums at least partially encompasses the effect of ability (in the form of state olympiads). A school receives the status of a state gymnasium for its great academic performance (Cabinet Regulations No. 129, 2001), and it is also required that the institution maintains the same level of education in the future. This, in turn, leads to the possibility that state gymnasiums can better prepare participants of various olympiads. The status of a state gymnasium is less pronounced for humanities (see Appendix 2). The effect on educational outcomes at gymnasiums is also positive and statistically significant at all specifications, only the actual impact is smaller and it is rather stable across different subjects.

Even though we should evaluate results on magnet and boarding schools with caution due to the small sample size, patterns of worse academic performance in these schools can be observed. The coefficients achieve statistical significance and take larger absolute values only when we do not account for student ability (state olympiad results are excluded). For instance, in regression (4) students from these institutions lose 6.433 points compared to secondary school seniors.

It is estimated that schools with Russian as the main language of instruction tend to have lower average exam grades than schools where all the subjects are taught in Latvian. The average loss of 3.168 points (regression (2)) can be associated with the fact that all the exam content is only in Latvian and some of the students can face certain difficulties while solving the exercises. Students might suffer from the same challenge also during the state olympiads and that shapes the 1% significance for the coefficient. The participation of Russian-speaking students in olympiads can be questioned as well; therefore, we believe that the importance of the coefficient is better described in regressions (2) and (4). The coefficients on mixed language schools are of different signs and are statistically insignificant for both average and humanities grades (see Table 1 and Appendix 2). If we look at the natural sciences exam grades, insignificant but constantly negative impact on grades is estimated. Even though the exam results of schools with other languages of instruction rather than Latvian and Russian statistically differs from schools with Latvian as the language of instruction at 1%, we treat this coefficient with caution as the sample again is small.

The regression coefficients for the natural science exam grades contradict the effects on the overall performance discussed previously (see Appendix 2). Schools with Russian language of instruction perform by on average 5.435 points better than schools with Latvian as the language of instruction, holding everything else constant (regression (5)). The statistical significance of this positive effect is just slightly reduced by the effect of the state olympiad variable (from 1% to 5%). An opposite effect is present when comparing the humanities exam grades among schools with Russian and Latvian as the languages of instruction. We find strong negative effect on study outcomes for schools with Russian as the language of instruction at 1% significance (see Appendix 2). Having Latvian as the language of instruction shows an average 6.79 point gain in the exam, holding other variables constant. This result is more in line with findings from the overall average exam grades; thus, we should expect the humanities exam grades to have greater weights in the overall performance parameters.

There is less space for discussion when describing dummy variables reflecting the location of the educational institutions. It can be seen in regressions (2), (3) and (4) that the coefficient is always negative, meaning that schools in Riga, on average, show better performance in standardized exams. Statistical significance of this phenomenon appears in the later regressions when the location dummies capture the ability effect that it otherwise suspected to be included in the state olympiad variable. The same argumentation is applicable in the case where the humanities exam grades are analyzed (see Appendix2). Fifth (5) specification portrays a rather controversial situation, as the value for city and rural schools is positive and rather large in absolute terms, but the value does not carry any statistical significance. Thus, we see the results from other regressions to have a better explanation on the actual effect of the urbanization dummies.

The regression (2) for average exam grades differs from others also in the respect of overall importance of obtaining secondary education in certain location. The F-test with value that equals 5.13 and p-value of less than 0% reveals that the average exam results should deviate in statistically significant manner across the country, when not accounting for the controls already elaborated upon. Specification (4) shows some signs of statistical discrepancies based on urbanization. The situation changes when state olympiad variable is introduced in regression. Average grades for natural science subjects (see Appendix 2) interdict the importance of obtaining education in certain location. Dummy variables on urbanization can be treated as statistically insignificant and the slightly worse results among schools in Riga can display only occasional discrepancies. When interaction terms are introduced for humanities, the importance of school location can again be questioned.

The adjusted R-squared of above discussed regression specification ranges from 26.6% to maximum of 31.3% for the average of all standardized exam grades. The introduced variables explain couple of percentage points more for humanities. Overall, we have reached higher R-squared values than in the studies cited in Woessmann (2000), as well as the actual estimation performed in Woessmann (2000). The R-squared values reported in the aforementioned source range from 17% to 25%. Here it is worth considering that the adjusted R-square is always smaller or equal to the unadjusted measure. Therefore, our regression specifications are able to explain more variance in the dependant variable than the mentioned previous studies.

6 Discussion

With this section we would like to combine the results from our model estimates with the qualitative insights we gained during the interviews. By comparing prior evidence in the literature with our empirical findings we aim to validate our findings as well as offer an explanation for them.

The coefficient on the pupil-teacher ratio retains its significance across all of the regression specifications employed in our study. Furthermore, it positively influences the educational outcomes in humanities, natural sciences and on average. This contradicts the negative observed effect on science scores in Ammermueller et al.(2003). While the estimated effects of pupil-teacher ratio according to review of studies in Hanushek (2006) are as likely to be positive, negative or insignificant, we find such consistence across specifications to be an indicator of real-life significance as well. Also M. Bērziņš (personal communication, March 31, 2011) mentions larger classes to have a better effect on educational outcomes due to the potential group dynamics and larger degree of rivalry within the group, which leads to a spiral effect.

In our starting regression specifications we find that the level of urbanization of a school's location has a positive and significant influence on the educational outcomes. When we start to control for ability we see the magnitude of significance to decrease. A reasonable explanation of this we see in the possibility that the more talented students gravitate towards higher levels of urbanization, as larger cities, larger areas have more possibilities outside the school for students to enrich themselves (V. Ņefjodova, personal communication, March 25, 2011). As the schools in areas with higher level of urbanization have better educational outcomes there is also the practice of introducing admission criteria, thus creating a barrier for transcending the levels of urbanization for low-performing students. Another possible explanation for this gravitation phenomenon could also come from the hunch that better teachers tend to gravitate towards higher level of urbanization in the same fashion as the students do. "There are not any great incentives for teachers to teach in rural areas" (V. Ņefjodova, personal communication, March 25, 2011). Although we do not have an extensive set of teacher characteristics included in our model specification, R. Kalvāns (personal communication, March 22, 2011) argues based on his professional observations that the teacher quality does not vary across levels of urbanization. Therefore, the gravitation of students seem to have more grounding than the gravitation of teachers. Still, this effect of controlling for ability when looking at urbanization is not observable when considering the educational outcomes in natural sciences.

Another approach on explaining the differing results across levels of urbanization is to examine the pupil-teacher ratio effects and how they interact with the urbanization variables. The interaction terms are included in our regression specification due to the fact that schools in rural areas in comparison with schools in urban areas have a lower required pupil-teacher ratio (Cabinet Regulations No. 1616, 2009). Due to this we suspect that the coefficient on the pupil-teacher alone might be too “averaged out”. As we see in the regression output (see Table 1 and Tables 5 & 6 in the Appendix 2) the effect of pupil-teacher ratio does significantly differ across urbanization levels when looking at the grade in humanities. As the introduction of these interaction terms takes away the significance of the urbanization dummies in both the specifications with and without the ability proxy, we see this as another explanation for the difference in grades among the levels of urbanization. Also we believe the group dynamics argument by M. Bērziņš (personal communication, March 31, 2011) should be of bigger importance for humanities, which we see in the higher statistical significance in the respective regression specification.

The evidence provided by the coefficients on the dummy variable accounting for schools with Russian as the language of instruction facilitates a discussion on both the ability and offered study conditions. Contradicting the findings of Geske et al. (2003), where based on TIMSS the authors conclude that there are no strong discrepancies between schools with Latvian or minority language as language of instruction, we find statistically significant disparities in exam grades between such schools. Still, it needs to be mentioned that Geske et al. (2003) control for family characteristics when making their claim.

Schools with Russian language of instruction perform on average worse than schools with Latvian as language of instruction in history and language exams, while they are more successful in mathematics, physics, chemistry and biology. We see several possible explanations for this observation.

One of the possible reasons why schools with Russian as the language of instruction show on average better results in natural science exams is the fact that these subjects demand less proficiency in Latvian language than e.g. humanities and that these two broad subject areas are perceived differently across ethnic groups. It was recognized during the interviews that many teachers have insufficient command of Latvian language that negatively affects the education quality (V.Ņefjodova, personal communication, March 25, 2011). It was also discussed in a research prepared by *Baltic Institute of Social Sciences* (2009) that previously special Latvian language courses for teachers were offered while recently such practice was seized. Also, V. Ņefjodova (personal communication, March 25, 2011) mentions that families with Russian ethnic background have a stronger "tradition" in natural sciences i.e. these sciences are perceived as more important.

Students face the same challenges. Their standardized exam results can be affected by the knowledge of language in which the exam problems are presented, namely Latvian language. Yet, comparisons of exams written by students from schools with Russian language of instruction reveal that there are no statistically significant differences in results irrespective of whether the exam was written in Latvian or Russian (R.Kalvāns personal communication, March 22, 2011).

An explanation for differences in history exams among schools can include some political insights as the perception about historic development of Latvia among different ethnicities can differ (R.Kalvāns, personal communication, March 22, 2011). If we look at the exam score data in history, we do see that the schools with Russian as the language of instruction do perform worse in the history exam than the schools with Latvian as the language of instruction. Yet, we cannot test in our data whether this is truly attributable to different attitudes towards history.

Also the family effects cannot be examined in our model, but the information from the interviews at least partially is in line with the evidence from Herderson & Mapp (2002) that parent involvement has positive effect on student performance. Our model specification assumes that those family effects are equal for all the schools. Still, the interviewees have observed that parents from schools with Russian language of instruction tend to have a higher involvement in the study process of students (V.Ņefjodova, personal communication, March 22, 2011). From this an argument arises that family effects are different among ethnic groups. On the contrary, our interviews also provided opinions where this involvement seems to be inconclusive among the different ethnic groups (M. Bērziņš, personal communication, March 31, 2011). The variation of family effects among ethnic groups has also been pointed out in Henderson & Mapp (2002). Thus, we come to the conclusion that the language of instruction might be picking up family effects.

Employing both languages of instruction (Latvian and Russian) does not reveal any significant dissimilarity in the overall exam grades contrasted to schools where the language of instruction is either Russian or Latvian. It was stated in an interview that evening (shift) schools that usually have both languages of instruction should represent lower standardized exam grades (R.Kalvāns, personal communication, March 22, 2011). However, as we do not have any evening (shift) schools included in the sample, we are not able to draw explanations why differences between Latvian and mixed languages of instruction could appear.

Schools with minority language of instruction display significantly better performance across both natural science and humanities exams compared to school with Latvian as language of instruction. Even though the sample for these institutions is small, we still observe statistically highly significant differences; therefore, we allow ourselves to make some conclusions regarding this phenomenon. During the interview it was unveiled that minority students might have higher motivation than non-minority students due to their perception of their position in society (V.Ņefjodova, personal communication, March 25, 2011). The line of argument showed that minority students might feel that they have to prove themselves to feel integrated in the general society.

The evidence from our model shows that state gymnasium results are indeed superior. Apart from the fact that state gymnasiums receive their status for the performance shown in the past, we would argue that this result is sustained due to the following two factors. First, the additional 10% monetary compensation as well as the expectations that the school will manage to attract better students might cause higher interest among most recognized teachers. As a consequence, students with highest academic potential will be interested in entering state gymnasiums because they feel confident that the education of best quality will be provided.

Similarly, we find that gymnasiums perform on average better than secondary schools. This could be explained by the reason that many gymnasiums have entry requirements and that makes an impression that higher level of education is provided (M.Bērziņš, personal communication, March 31, 2011). Also the brand aspect of the names of both gymnasiums and state gymnasiums can affect the choice of entering a particular educational institution (R.Kalvāns, personal communication, March 22, 2011), but it does not necessarily mean that the provided education will always respond to the high expectations. Comparing the results across the country, *State Curriculum Centre* representatives have noticed that there are several gymnasiums constantly receiving lower average grades relative to secondary schools (R.Kalvāns, personal communication, March 22, 2011).

Boarding and magnet schools achieve lower results than any other type of educational institution presented in our regression. The mean exam grade value for boarding and magnet schools was 43.4 for the second observation year, whereas the corresponding value for secondary schools was 50.6. Gymnasiums together with state gymnasiums have even higher standardized exam scores (see Appendix 1, table 4). We might argue that the lower standardized exam grade is the result of boarding and magnet schools serving a different function than the more conventional establishments of general education.

Statistically insignificant effect of teacher age on academic achievements can be explained by the reason that soft skills matter more than a certain level of education or age related experience (R.Kalvāns, personal communication, March 22, 2011; V.Ņefjodova, personal communication, March 25, 2011; M.Bērziņš, personal communication, March 31, 2011). This opposes the finding by Woessmann (2003c) that age is a strong negative determinant of student performance. Also we should take into account that the average teacher age does not exhibit a large degree of variation in Latvia (see Appendix 1). Due to this econometric methods can show only limited results.

7 Conclusions

With this study we have examined the effect of school-specific factors on educational outcomes. To accomplish this we have created a unique dataset on school-specific factors and educational outcomes by combining three different data sources both from public authorities and a private institution. At our present state of knowledge such a dataset has not been created previously in Latvia. Also by employing the model of education production function we have gained a deeper insight in the distribution of educational outcomes and most importantly the reasons for it among different secondary schools in Latvia. Furthermore, we are not aware of other estimates of the education production function for Latvia, which employ the school level educational outcomes.

At the beginning of this paper we set ourselves an aim to search for explanations for the diverse performance of schools in Latvia. First, we find that the type of a school and the language of instruction at the school significantly segment the landscape of schools in Latvia across the different specifications of educational outcomes, while the level of urbanization of a school’s location significantly divides the performance of schools in terms of overall performance and the performance in humanities. Second, we find that these categorical differences can be attributed to the positive and significant effect of the pupil-teacher ratio and its interaction with the level of urbanization and the positive and significant effect of ability. We also develop the discussion on the underlying mechanics of these uncovered relations by employing qualitative insights from experts within the field of education in Latvia via methodological triangulation.

The findings of this research point out the schools which are likely to have worse educational outcomes than average and also show the possible reasons for it. Also the results from segmentation of schools by the school-specific factors and their effect on the school performance have policy implications in the form of which schools should be targeted for support programs. Similarly, the consistently positive and significant effect of the pupil-teacher ratio shows support to the policy of increasing the minimum pupil-teacher ratio in all schools regardless of the level of urbanization as a cost-effective way of improving educational outcomes.

Although we managed to include a very high share of schools which offer secondary education, this is also our pitfall in extending the dataset to include more factors due to the inherent difficulties in gathering the data for the same entities on such a large scale. In the case where we could obtain additional data, we would like to relax our assumption in the model on the family and peer characteristics by accounting for them directly. Also, although expert opinions argued for the irrelevance of the education of teachers (beyond the minimum), we still would have preferred to test it empirically. While we did not observe time fixed effects to be significant, we see an extension of the time span of the dataset as a welcome addition to a further study.

References

Amendment of Education Law. (2004). Retrieved January 22, 2011, from http://www.likumi.lv/doc.php?id=84315

Ammermueller, A., Heijke, H., & Woessmann, L. (2003). *Schooling Quality in Eastern Europe: Educational Production during Transition.* (IZA Discussion Paper Series IZA DP No. 746).

Andrews, D.R., Fayissa, B., & Tate, U.S. (1991). An Estimation of Aggregate Educational Production Function for Public Schools in Louisiana. *The Review of Black Political Economy 20*(1), 25-47.

Baltic Institute of Social Science. (2009). *Centralizēto eksāmenu pēdējo trīs gadu rezultātu analīze sagatavotības pārejai uz vienotu latviešu valodas eksāmenu 2012.gadā izpēte.* Retrieved April 1, 2011, from http://izm.izm.gov.lv/upload_file/Registri_statistika/IZM-petijums-pareja-uz-vienotu-latv-val-eksamenu.pdf

Boissiere, M. (2004). *Determinants of Primary Education Outcomes in Developing Countries.* (Background Paper for the Evaluation of the World Bank's Support to Primary Education). Washington, D. C.: The World Bank.

Bryman, A., & Bell, E. (2003). *Business Research Methods*. Oxford: Oxford University Press.

Buddin, R., & Zamarro, G. (2009). *Teacher Qualifications and Middle School Student Achievement.* (Institute of Education Sciences Working Paper No. WR-671-IES).

Cabinet Regulations No. 129. (2001). *Ģimnāzijas un valsts ģimnāzijas statusa piešķiršanas un anulēšanas kārtība un kritēriji*. Retrieved January 21, 2011, from http://www.likumi.lv/doc.php?id=5771

Cabinet Regulations No. 347. (2000). *Noteikumi par prasībām pegadogiem nepieciešamajai izglītībai un profeionālajai kvalifikācijai*. Retrieved February 18, 2011, from http://www.likumi.lv/doc.php?id=11488

Cabinet Regulations No. 1616. (2009). *Kārtība, kādā aprēķina un sadala valsts budžeta mērķdotāciju pašvaldību un privātajām izglītības iestādēm bērnu no piecu gadu vecuma izglītošanā nodarbināto pirmsskolas izglītības pedagogu darba samaksai un pašvaldību vispārējās pamatizglītības un vispārējās vidējās izglītības* *iestāžu pedagogu darba samaksai.* Retrieved November 17, 2010, from http://www.likumi.lv/doc.php?id=202915

Coleman, J.S. (1966). *Equality of Educational Opportunity (Coleman) Study (EEOS).* Washington, DC: Office of Education.

The Constitution of the Republic of Latvia. (2010). Retrieved November, 20, 2010, from http://www.likumi.lv/doc.php?id=57980

Draudzīgā aicinājuma fonds. (2009). *Datu bāzē reģistrētās olimpiādes un konkursi. 2008./2009. mācību gads.* Retrieved March 1, 2009, from http://www.konkurss.lv/r_da_balva_not.asp?Mac_g=2008./2009.

Draudzīgā aicinājuma fonds. (n.d.). *Draudzīgā aicinājums fonda vēsture*. Retrieved February 18, 2011, from http://www.konkurss.lv/da_fonds/da_fonds.asp

Dreijere, V., & Valaine, I. (2009). *Lai Nauda Neaizietu Neceļos.* Retrieved November 17, 2010, from http://diena.lv/lat/laikraksts/latvija/lai-nauda-neaizietu-necelos

Ederer, P., Schuler, P., & Willms, S. (2007). *The European Human Capital Index: The Challenge of Central and Eastern Europe.* Retrieved January 22, 2011, from http://www.ibm.com/ibm/governmentalprograms/lisbon_council_european_human_capital_index_cee.pdf

Education Law. (2011). Retrieved January 17, 2010, from http://www.likumi.lv/doc.php?id=50759#4582

Eurostat. (2010). *Pupil/Student - Teacher Ratio and Average Class Size (ISCED 1-3).* Retrieved December 17, 2010 from http://epp.eurostat.ec.europa.eu/

Geske, A., Grinfelds, A., Dedze, I., & Zhang, Y. (2006). School Quality and Equity in Central and Eastern Europe. Family Background, School-Quality and Rural-Urban Disparities in Student Learning Achievement in Latvia. *Prospects 36*(4), 419-431.

Geske, A., Grinfelds, A., & Kangro, A. (2003). *International Comparative Educational Research in Latvia: Current Results and Trends.* Retrieved January 21, 2011, from http://www.leidykla.vu.lt/inetleid/acta_pae/10/straipsniai/str7.pdf

Hanushek, E.A. (2006). School Resources. In E.A. Hanushek, & F. Welch (Eds.) *Handbook on the Economics of Education, Vol. 2* (chap. 14). Amsterdam: North Holland.

Hanushek, E.A. (2007). Education Production Functions. In S.N. Durlauf, L.E. Blume (Eds.) *The New Palgrave Dictionary of Economics, 2nd ed*. New York: Palgrave Macmillan.

Hanushek, E.A., Rivkin, S.G., & Taylor, L.L. (1996). Aggregation and the Estimated Effects of School Resources. *The Review of Economics and Statistics 78*(4), 611-627.

Henderson, A.T., & Mapp, K.L. (2002). *A New Wave of Evidence. The Impact of School, Family, and Community Connections on Student Achievement.* Austin, TX: Southwest Educational Development Laboratory.

Iespējamā misija. (n.d.). *Ideja*. Retrieved February 16, 2011, from http://www.iespejamamisija.lv/lv/par_mums/ideja_un_vizija

Jankelevica, M. (2010). *"Draudzīgā aicinājuma" balvu pasniegs pēdējo reizi*. Retrieved February 19, 2011, from http://zinas.nra.lv/latvija/31567-draudziga-aicinajuma-balvu-pasniedz-pedejo-reizi.htm

Kuzmina, I. (2010). *Eksperti iebilst pret skološanos tikai latviski*. Retrieved January 22, 2011, from http://www2.la.lv/lat/latvijas_avize/la_pielikumi/latvijas.zinas/?doc=71754

Law on Administrative Territories and Populated Areas (Administratīvo teritoriju un apdzīvoto vietu likums). (2011). Retrieved January 31, 2011, from http://www.likumi.lv/doc.php?id=185993

Levacic, R., Jenkins, A., Vignoles, A., & Allen, R. (2005). *The Effect of School Resources on Student Attainment in English Secondary Schools.* Retrieved November 22, 2010, from http://www.cepr.org/meets/wkcn/3/3525/papers/Vignoles2.pdf

Mathison, S. (1988). Why Triangulate? *Educational Researcher 17(*2*)*, 13-17.

The Ministry of Education and Science. (2009). *Izglītības funkciju nodrošināšana pēc administratīvi teritoriālās reformas un finansēšanas modeļa "Nauda Seko Skolēnam" ieviešanas. Ieteikumi pašvaldībām un izglītības darba vadītājiem.* Retrieved November 17, 2010, from http://izm.izm.gov.lv/upload_file/rokasgramata-nauda-seko-skolenam-001.pdf

NIID.LV (Nacionālā Izglītības Iespēju Datu Bāze). (2008). *Valsts ieskaites un centralizētie eksāmeni 2008./2009. mācību gadā*. Retrieved November 19, 2010, from http://www.niid.lv/node/330

NIID.LV (Nacionālā Izglītības Iespēju Datu Bāze). (2010). *Pirmo svešvaloda sāks mācīt no 1.klases*. Retrieved February 20, 2011, from http://www.niid.lv/lv/node/547

State Curriculum Centre (Valsts Izglītības Satura Centrs). (n.d.). *Eksāmenu kārtotāju skaits no absolventu skaita vispārizglītojošās skolās*. Retrieved November 19, 2010, from http://visc.gov.lv/eksameni/vispizgl/statistika/2010/dokumenti/!visi1_eks_nek.jpg

Vincevs, A. (2006). *Atbirst, jo nezina eksaktos*. Retrieved February 18, 2011, from http://www.apollo.lv/portal/life/articles/88931

Woessmann, L. (2003a). *Educational Production in East Asia: The Impact of Family Background and School Policies on Student Performance.* (IZA Discussion Paper Series IZA DP No. 745).

Woessmann, L. (2003b). *European "Education Production Function": What Makes a Difference for Student Achievement in Europe?* (European Economy. Economic Paper No. 190/2003).

Woessmann, L. (2003c). Schooling Resources, Educational Institutions and Student Performance: the International Evidence. *Oxford Bulletin of Economics and Statistics 65*(2), 117-170.

Appendix 1 – Variables and sample statistics

Definitions of variables

Variable	Definition	Type
Averagegrade	Average grade of all written exams	Numeric
Naturalgrade	Average grade of all exams written in Physics, Chemistry, Math and Biology	Numeric
Humgrade	Average grade of all exams written in History and languages (State and foreign languages)	Numeric
Teacherage	Average age of all teachers employed in specific school	Numeric
Pupilteacher	Pupil-teacher ratio	Numeric
g2	Secondary school	Binary
g3	Gymnasium	Binary
g4	State gymnasium	Binary
g5	Magnet and boarding school	Binary
valoda1	School with Latvian as instruction language	Binary
valoda2	School with Russian as instruction language	Binary
valoda3	School with mixed (both Latvian and Russian) languages as instruction language	Binary
valoda4	School with minority (Lithuanian, Ukrainian etc.) language as instruction language	Binary
urban1	School situated in Riga	Binary
urban2	School located in Republican city	Binary
urban3	School situated in small city	Binary
urban4	School located in rural areas	Binary

Table 2 – Definitions of variables. Source: made by authors

Summary statistics of numeric variables

Variable	Year	Mean	St.Dev.	n
Averagegrade	2008	50.63	9.22	233
	2009	52.11	9.34	343
	All	51.51	9.31	576
Naturalgrade	2008	45.22	11.51	233
	2009	46.08	10.96	343
	All	45.73	11.18	576
Humgrade	2008	53.06	9.88	233
	2009	54.93	9.99	343
	All	54.17	9.98	576
Teacherage	2008	45.82	2.87	233
	2009	45.06	2.76	343
	All	45.37	2.83	576
Pupilteacher	2008	9.89	2.73	233
	2009	12.24	7.10	343
	All	11.29	5.86	576

Table 3 – Summary statistics of numeric variables. Source: made by authors

Note: 2008 and 2009 refer to the academic years 2008/2009 and 2009/2010 respectively.

Summary statistics of numeric variables across categories

Binary variable	Year	n	Averagegrade		Naturalgrade		Humgrade		Pupilteacher	
			Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
Secondary School	2008	188	49.1	8.71	43.6	11.00	51.6	9.66	9.86	2.59
	2009	286	50.6	8.84	44.6	10.33	53.5	9.76	11.80	5.75
Gymnasium	2008	23	56.1	6.62	48.7	8.49	59.3	6.94	9.95	3.60
	2009	30	58.4	6.54	51.1	9.42	61.8	6.18	16.00	14.59
State Gymnasium	2008	17	62.2	5.13	60.7	4.32	63.1	6.76	11.20	2.29
	2009	22	64.5	4.95	61.0	6.40	66.2	5.18	13.60	7.07
Boarding & Magnet Sch.	2008	5	43.1	10.40	36.1	15.76	47.0	7.39	6.29	1.52
	2009	5	43.4	7.87	34.6	7.30	47.5	8.17	9.90	3.99
Latvian language	2008	175	51.1	8.66	44.9	11.10	53.9	9.01	9.99	2.55
	2009	229	52.9	8.73	45.6	11.21	56.3	8.73	12.30	7.49
Russian language	2008	36	48.4	11.50	48.0	11.69	48.5	13.27	9.66	3.46
	2009	86	50.7	10.85	48.6	10.28	51.8	12.25	12.80	6.58
Mixed language	2008	20	49.2	8.60	41.9	13.74	52.4	7.77	9.63	2.98
	2009	25	49.3	8.63	41.7	9.14	52.4	9.82	11.20	4.97
Minority language	2008	2	64.1	4.04	57.9	8.99	66.7	1.86	7.91	0.22
	2009	3	59.3	3.89	50.2	10.78	61.8	7.53	5.49	1.82
Riga	2008	43	53.6	11.42	47.1	12.05	56.3	12.66	10.40	3.53
	2009	98	54.5	9.96	47.2	11.30	57.6	10.59	14.80	10.79
Republican city	2008	30	53.4	8.79	48.2	10.61	55.8	8.81	11.00	2.34
	2009	59	53.4	9.85	49.2	10.71	55.4	10.80	13.60	5.79
Small city	2008	63	51.2	9.30	46.9	11.28	53.0	9.84	10.30	2.35
	2009	87	52.7	8.41	46.7	10.41	55.6	8.94	11.60	4.29
Rural area	2008	97	48.1	7.48	42.3	11.21	50.8	8.24	9.10	2.48
	2009	99	48.5	8.17	42.6	10.46	51.4	8.83	9.50	3.09

Table 4 - Summary statistics of numeric variables across categories. Source: made by authors

Note: 2008 and 2009 refer to the academic years 2008/2009 and 2009/2010 respectively.

Appendix 2 – Regression results

Dependent variable:	***Natural science exam grade***				
Regressor	(1)	(2)	(3)	(4)	(5)
Average teacher age	0.086 (0.158)	0.081 (0.160)	-0.126 (0.256)	0.088 (0.160)	-0.037 (0.249)
Pupil - teacher ratio	0.448*** (0.081)	0.451*** (0.084)	0.734** (0.319)	0.454*** (0.107)	1.252** (0.587)
Gymnasium	5.207*** (3.648)	5.143*** (1.326)	4.580** (1.785)	5.097*** (1.344)	4.576*** (1.716)
State gymnasium	16.800*** (0.998)	16.490*** (1.139)	14.589*** (1.713)	16.487*** (1.149)	14.488*** (1.764)
Boarding and magnet school	-6.312* (3.648)	-6.160* (3.685)	-2.944 (6.930)	-6.287* (3.692)	-2.544 (7.431)
Russian language	4.535*** (1.079)	4.621*** (1.226)	5.624** (2.241)	4.593*** (1.230)	5.435** (2.223)
Mixed language	-1.363 (1.630)	-1.549 (1.640)	-0.714 (2.967)	-1.495 (1.622)	-0.108 (2.941)
Minority language	8.940*** (3.258)	9.212*** (3.372)	17.561*** (5.309)	9.230*** (3.394)	18.754*** (5.604)
State city	-	0.580 (1.279)	0.132 (2.331)	0.548 (3.062)	2.419 (13.241)
City	-	0.734 (1.242)	1.412 (2.278)	0.293 (3.010)	4.106 (8.707)
Rural area	-	0.177 (1.360)	1.180 (2.504)	1.063 (3.349)	12.258 (8.391)
Pupil-teacher ratio × State city	-	-	-	0.002 (0.215)	-0.236 (1.131)
Pupil-teacher ratio × City	-	-	-	0.039 (0.229)	-0.265 (0.779)
Pupil-teacher ratio × Rural area	-	-	-	-0.096 (0.296)	-1.152 (0.771)
State Olympiad results	-	-	0.683*** (0.166)	-	0.656*** (0.168)
Intercept	34.324*** (7.385)	34.216*** (7.277)	39.732*** (11.584)	33.869*** (7.282)	30.321** (12.590)
F-statistics testing the hypothesis that the population coefficients on the indicated regressors are all zero:					
School type	96.92 (0.00)	72.85 (0.00)	24.51 (0.00)	71.46 (0.00)	22.91 (0.00)
School location	-	0.16 (0.92)	0.23 (0.88)	0.04 (0.99)	0.85 (0.47)
Language of instruction	8.92 (0.00)	7.01 (0.00)	4.78 (0.00)	6.91 (0.00)	4.94 (0.00)
Pupil-teacher ratio × School location	-	-	-	0.05 (0.98)	0.88 (0.45)
Regression Summary Statistics					
R^2	0.258	0.286	0.256	0.259	0.269
$\overline{R}^2$	0.246	0.243	0.216	0.239	0.218
n	576	576	233	576	233

Table 5 – Regression results using *naturalgrade* as the dependant variable. Source: made by authors

Note: All regressions include an intercept. Heteroskedasticity-robust standard errors are given in parenthesis under estimated coefficients. *p*-values are given in parentheses under *F*-statistics, which is heteroskedasticity-robust. Coefficients are significant at the *10%, **5%, ***1% significance level.

Dependent variable:	*Humanities exam grade*				
Regressor	(1)	(2)	(3)	(4)	(5)
Average teacher age	0.197 (0.133)	0.255* (0.132)	0.244 (0.198)	0.236* (0.132)	0.359* (0.192)
Pupil - teacher ratio	0.460*** (0.092)	0.367*** (0.088)	0.982*** (0.224)	0.323*** (0.107)	1.739*** (0.371)
Gymnasium	6.850*** (0.990)	5.556*** (1.018)	6.377*** (1.360)	5.706*** (0.993)	6.648*** (1.332)
State gymnasium	10.271*** (1.052)	9.915*** (1.123)	7.576*** (2.151)	9.985*** (1.147)	7.619*** (2.055)
Boarding and magnet school	-4.767** (2.332)	-6.271** (2.903)	-2.856 (3.741)	-5.888* (3.006)	-1.892 (3.330)
Russian language	-3.940*** (1.140)	-6.455*** (1.282)	-6.790*** (2.304)	-6.351*** (1.272)	-7.057*** (2.313)
Mixed language	-1.255 (1.259)	-1.535 (1.471)	0.475 (2.014)	-1.645 (1.489)	0.801 (2.184)
Minority language	10.471*** (3.913	6.204* (3.560)	11.949*** (2.681)	6.052* (3.498)	13.778*** (2.799)
State city	-	-3.179** (1.324)	-3.191 (2.194)	-4.070 (2.919)	0.653 (7.176)
City	-	-5.126*** (1.239)	-5.606** (2.204)	-5.577* (3.032)	7.862 (8.101)
Rural area	-	-5.941*** (1.269)	-4.301* (2.221)	-9.525*** (2.545)	7.068 (5.262)
Pupil-teacher ratio × State city	-	-	-	0.069 (0.206)	-0.385 (0.619)
Pupil-teacher ratio × City	-	-	-	0.037 (0.226)	-1.307* (0.743)
Pupil-teacher ratio × Rural area	-	-	-	0.373* (0.203)	-1.144** (0.461)
State Olympiad results	-	-	0.624*** (0.175)	-	0.597*** (0.180)
Intercept	39.638*** (6.223)	42.696*** (6.248)	32.158*** (10.283)	44.047*** (6.230)	21.896** (10.048)
F-statistics testing the hypothesis that the population coefficients on the indicated regressors are all zero:					
School type	46.27 (0.00)	32.60 (0.00)	10.14 (0.00)	32.24 (0.00)	11.30 (0.00)
School location	-	7.96 (0.00)	2.19 (0.09)	4.72 (0.00)	0.83 (0.48)
Language of instruction	6.97 (0.00)	11.07 (0.00)	17.29 (0.00)	10.96 (0.00)	18.46 (0.00)
Pupil-teacher ratio ×School location	-	-	-	1.16 (0.32)	2.48 (0.06)
Regression Summary Statistics					
R^2	0.252	0.292	0.340	0.295	0.361
$\overline{R}^2$	0.240	0.277	0.304	0.276	0.317
n	576	576	233	576	233

Table 6 - Regression results using *humgrade* as the dependant variable. Source: made by authors

Note: All regressions include an intercept. Heteroskedasticity-robust standard errors are given in parenthesis under estimated coefficients. *p*-values are given in parentheses under *F*-statistics, which is heteroskedasticity-robust. Coefficients are significant at the *10%, **5%, ***1% significance level.

Appendix 3 – Interview Questions

Introduction: Thank you for agreeing to participate in this interview. At the moment we are at the final stage of our research and we have already performed our quantitative data analysis. Therefore, we would highly appreciate your expert input by providing your experiences and knowledge about the education system in Latvia to help us to develop a deeper insight in our findings.

1. How would you characterize in overall the education system and, in particular, secondary education in Latvia?
2. What to your mind are the main strengths and weaknesses of secondary education in Latvia?
3. Have you noticed any differences in academic performance of pupils during the last years?
4. What in your opinion has the greatest effect on the overall student results at secondary education level?
5. What other factors to your mind should be taken into account when analyzing student performance?
6. Do you think that there are any significant differences between gymnasiums and secondary schools, especially in relation to standardized exam results?
7. Have you noticed any significant discrepancies among Russian, Latvian and two-stream schools?
8. How big is the importance of teachers' competence and how would you evaluate it (depending on education, seniority and other factors)?
9. How could you explain the phenomenon that Russian students perform better in Natural Science subjects, while Latvians are superior in Humanities?

Note: these are only broad categories of questions which were asked during the interviews.