universität
wien

MAGISTERARBEIT

Titel der Magisterarbeit

„ Class Size Effects and the Educational Production Function in Austrian Schools "

Verfasser

Michael Topf, Bakk.rer.soc.oec.

angestrebter akademischer Grad

Magister der Sozial und Wirtschaftswissenschaften (Mag.rer.soc.oec)

Wien, 2012

Table of Contents

1. Introduction

A corner stone in development of modern society is education. Improvements in technology, science or other sections were mainly achieved through increased education. Education is also related to economic growth (Hanushek and Kimko, 2000). An extra year of education for men raises the growth rate by 1.2 % per year (Barro, 1997). A different point of view knowing that years of education contribute to economic growth, however, is to enhance educational quality from the microeconomic perspective. Jamison et al. (2006) found a link between educational quality and growth, stating that a standard deviation increase in test scores of students is associated with a 0.5 to 0.9 % growth in income per capita. Prior to fostering economic growth through improvement its educational quality one has to be aware of the factors that determine educational quality.

Although the educational level in Austria is certainly above average in recent years, discussions about educational reforms[1] and petitions for referendums[2] were induced by below-average performance at international student assessments[3]. Hence, concepts like the new secondary school[4] and comprehensive schools have been discussed in recent years. Until 2018 all lower secondary schools will be reformed into new secondary schools. Just this reform is reason enough to evaluate the determinants of students' success in Austrian schools, since further improvements of the system might coincide with such a big reform.

This paper analyzes the impact of student and family characteristics on student achievement. Furthermore the impact of differences between school types will be evaluated and discussed. The main focus, however, is to determine the effect of class size on student achievement in Austrian schools. By investigating two large scale studies namely the Programme for International Student Assessment study (PISA) and the Progress in International Reading Literature Study (PIRLS) that aim to measure individual ability of students within and across countries, the impact of class size on test scores will be investigated. There are numerous factors that might influence students'

---

1 http://www.diebildungsreform.at; http://www.bmukk.gv.at/medienpool/19400/bildungsreform.pdf last online: 24.04.2012

2 http://www.vbbi.at/

3 Schwantner & Schreiner (2010)

4 http://www.neuemittelschule.at/ last online: 23.04.2012.

success. The aggregate of the factors that determine student achievement can be seen as an educational production function (EPF). For decades scientists Rice (1902),Card & Krueger (1992), Hanushek (1994), Hanushek (1999), Hanushek (2003), Krueger (1999), Krueger (2003), Akerhielm (1995), Angrist & Lavy (1999), Hoxby (2000), Wößmann & West (2002), Wößmann (2003) have been studying the effects of the EPF on student achievement. A substantial part of this literature focuses on the particular class size question. This is particularly important, because of the great costs that come along with smaller classes.

A reduction in class size requires more classrooms, more teachers and more administrative staff – all of them generating costs for society. The question is whether these costs are smaller than the benefits that come along with smaller classes. These benefits could be represented by an increase in the educational level, which induces an increase in years of education (Jamison et. al, 2006). For the duration of this paper the benefits of smaller classes are defined by an increase in test scores, which are assumed to be a reasonable approximation of an individual's educational level. As Austrian researchers found out when dealing with PISA 2000 and PISA 2003 data, class composition does have an impact on student performance. Schneeweis & Winter-Ebmer (2007) state that, "*the peer group effect in reading achievement is positive and diminishing in socioeconomic background. Thus, students from disadvantaged socioeconomic backgrounds have a higher return from a favorable peer group*".[5]

The new secondary school reform might perfectly coincide with a restructuring and a refinement of the whole educational system. Especially in the education industry, itself being a cornerstone of modern society, these reforms should be elaborated thoroughly. Thus, an investigation of the Austrian educational production function is necessary. Reacting to the latest performance of Austrian pupils in International student assessments, the ministry of education developed a ten-point-program, whose main goal is to raise the standard level of education. Language skills should be mediated already in kindergarten, expansion of all-day schools with intensified support, providing two teachers for German, Math and English classes in the new secondary school and

---

5 Haider et al. (2001) provided the framework for the empirical analysis. Their analysis is very similar to the analysis applied in this paper. Instead of peer group effects, class size effects will be determined.

establishment of a feed-back-culture in schools to investigate the basic skills after the 4th and 8th grade. A first investigation[6] of basic skills was conducted in early 2011 for nearly 30.000 Viennese pupils. All participants received their own test score after the examination and aggregated results have been provided to the public. It was reported that nearly one-quarter of all pupils had bad reading skills.[7]

Another one of these ten points states that class size should be reduced further. According to previous literature it is not clear whether a class size reduction is beneficial for student outcome. Besides, larger classes could even increase educational quality if students benefit from each other (i.e. spillover effects). According to theory there is no clear prediction of the class size effect, hence econometrical studies have to yield guidance. Wößmann & West (2002) who analyzed 18 participating countries of the Trends in International Mathematics and Science study ruled out large effects in 11 countries (Belgium, Canada, Czech Rep., Korea, Portugal, Romania, Singapore, Slovenia, Spain). Only in two countries, namely Greece and Iceland, they found sizeable positive effects of smaller classes. In Japan and Singapore no effects were found. In Australia, Hong Kong, Scotland and the United States their results are imprecise. Their findings suggest that effects might differ between countries, simply because schooling systems are different all around the world. Wößmann & West (2002), did not analyze Austrian data. Hence, it is important to investigate whether class size effects are present in the Austrian system. Wößmann (2003a) used the International TIMSS Database where Austrian pupils represented a minor share in the analysis of a total of 39 countries. He found that centralized examinations, control mechanisms and school autonomy in personnel are major determinants of a successful schooling system. When pooling all countries he finds a positive link between class size and student achievement.

By investigating two major standardized tests, namely PISA and PIRLS, this paper estimates the magnitude of the Austrian educational production function determinants. Using a grade average class size instrumental variable technique as well as a regression

6 SOKO – Lesen; http://www.stadtschulrat.at/aktuell/ last online: 23.04.2012.

7http://www.stadtschulrat.at/files/content_dl_1/Zusammenfassende_Ergebnisse.pdflast online: 23.04.2012.

discontinuity design this paper tries to determine the class size coefficient. These instrumental variable strategies are necessary due to an endogeneity bias of class size. Schools might apply policies that select either better or worse students into particular classes. Hence, it might be that the population between classes is different. To control for that endogeneity instrumental variables have to be included. The identification strategy will be explained in more detail in the measurement framework section. In this paper mixed evidence on the class size coefficient was found. Additionally it was found that the various school types in Austria attract different student populations on the one hand and contribute to the differences in student achievement on the other hand. This paper will continue as follows. In the next section I will provide a short class size literature overview. The third section introduces the non-familiar reader to the Austrian schooling system and the fourth section will describe the datasets. In the fifth section the measurement framework will be described followed by the description of the results in section six. Section seven concludes.

2. Literature Overview

Many studies, Coleman (1966), Card & Krueger (1992), Akerhielm K. (1995), Bishop (1997), Hoxby (1998), Angrist & Lavy (1999), Hoxby (2000), Levin (2001) and Wößmann & West (2002), focusing on the effect of class size have been published in the past centuries.

The class size question within this broad literature has always been of particular interest to researchers. Moses Maimonide, a rabbinic scholar, dealt with this topic as early as the 12th century and suggested a maximum class size rule of 40 in Israeli schools after interpreting the Talmud. One of the first scientific papers (Rice, 1902) in this area found no link between smaller classes and higher scholastic achievement; at a time where the number of students in class was three-times higher than it is today. Coleman (1966) heated up the debate on school quality with one of the largest studies in history of the EPF, known as the Coleman Report. Coleman along with other researchers analyzed data of more than 150.000 students.

At a time of racial discrimination and segregation he found that a portion of the divergence in education of whites and blacks can be attributed to differences in school

quality. On the one hand blacks were mostly enrolled in schools where the share of minority students was between 50 – 100 %. On the other hand, differences in student achievement stem from variation in student characteristics, rather than differences in school funding; a rather surprising finding that has bothered researchers since then. If school funding, compared to family background such as socioeconomic status, is of minor importance, why would one want to increase public education expenditures? The litmus test is to find out the effects that matter and customize the education system according to the important criteria. Over time the "money makes no difference" finding has been investigated by many researchers, who found little evidence to challenge these landmark findings.

The class size literature, itself being one of the most investigated subtopics of the broad school quality literature, lacked of convincing evidence as well - until the Tennessee STAR Project was conducted. Project STAR - a large scale randomized experiment designed to measure the effects reduced class size on scholastic achievement - features prominently in the class size debate. The set up was simple. Nearly 12000 students have been selected randomly into three different class size categories starting in kindergarten until the third grade. The categories were: 1) small classes (13-17 pupils), 2) regular classes (22-26 pupils) and 3) regular classes with an additional teaching aid (22-26 pupils). They were created in 79 different participating schools. At the end of each year a Stanford Achievement Test (SAT) and a Tennessee Basic Skill Test was administered, measuring educational achievement of the pupils.

Folger & Breda (1989), Finn & Achilles (1990) and Word et al (1990) found that students in smaller classes performed significantly better than students in larger classes. Krueger (1999) reinvestigating the STAR data and taking into account potential drawbacks (attrition, re-randomization after kindergarten, nonrandom transitions, variability of class size) supported the preceding findings and also found that every additional year in smaller classes yields positive but decreasing benefits in terms of scholastic achievement. A popular criticism of project STAR is the so-called Hawthorne-effect introduced by Landsberger (1958), which states that participating students know to be observed and evaluated, hence increase their effort.

Another argument states, that participating *"…schools and teachers could have anticipated that the outcome of the experiment is pioneering for future school funding policies. Thus, they might have put more effort into smaller classes to manipulate the outcome of the experiment"* Hoxby (2000).

In an earlier paper, using natural population variation, Hoxby (1998) neither found effects of smaller classes, nor could she link the presence of black students in a class with significant differences in student achievement. Exploiting a maximum class size rule of forty in Israeli schools, known as Maimonides' rule[8], Angrist & Lavy (1999) using a regression discontinuity design find a positive association between smaller classes and math and reading performance for fifth graders. The results for fourth graders were not statistically significant.

Why increasing the effectiveness of the education system and what are the consequences of it? Card & Krueger (1992) found that an increase in school quality in the United States coincides with an increase in average earnings of students. Another finding is that a decrease of the pupil/teacher ratio (which is an approximation for class size) by 10 students raises average education by 0.6 years, hence increase expected earnings by 3.2 %. In one of the more recent papers Denny & Oppedisano (2010), who analyze UK and US PISA 2003 data on mathematics and science scores, find significant positive effects of larger classes on student achievement. Additionally an experimental instrumental variable technique developed by Lewbel (2010) was applied in Denny & Oppedisano (2010). Using grade average class size and the experimental approach as an instrument they estimate a positive class size coefficient of 8.24 for the UK and 2.11 for the United States in mathematics.[9]

The next section will give a quick overview over the Austrian schooling system. Since one major finding of this paper is, that the structure of the Austrian schooling system does have a strong impact on student performance, it is important to review it.

---

[8] *'The number of pupils assigned to each teacher is twenty-five. If there are fifty, we appoint two teachers. If there are forty, we appoint an assistant, at the expense of the town''* (quote from Chapter II, page 21:a of the Baba Bathra; English translation on page 214 of Epstein [1976]).

[9] The UK coefficient is significant at the 5 % level and the US coefficient is not significant at the 10 % level.

3. The Austrian Schooling System

In Austria children typically start their schooling career at the age of 3, namely in kindergarten (ISCED 0).[10] At the age of 6 the students switch to primary school (ISCED 1). If students do not satisfy the requirements of the first grade when switching to school then they might be downgraded to a preschool; but only for one year. Primary school lasts four years. Then the students have the option of either attending a lower secondary school (APS) or academic secondary school (AHS). This, however, depends on the children's grades in primary school (both ISCED 2). Already in an early stage of a student's career, ability sorting takes places in the Austrian system. Academic secondary school is being considered as the type of school that prepares pupils for higher education. The general belief is that lower secondary schools prepare students for apprenticeships after they completed compulsory schooling, which is 9 years in Austria.

Lower secondary schools last four years, whereas academic secondary schools typically last eight years. In academic secondary schools, after four years, students have the option of either staying at school and to go for their secondary school leaving certificate called "Matura", or leaving school. Students that go to a lower secondary school could, depending on their interest, either attend a prevocational school (ISCED 3C – one year to complete compulsory schooling), an intermediate technical and vocational school (BMS – ISCED 3B) or a higher technical and vocational college (BHS – ISCED3A/4A), but only if the grades are sufficient. The students attending a prevocational school, have the chance to attend a vocational school for apprentices (BS – ISCED 3B) which is nearly equatable to intermediate technical and vocational schools. These school types typically focus on imparting special on the job knowledge (e.g. special knowledge for electricians or motorcar mechanic). A student leaving school with a "Matura" has the option of going to university. Concerning university education, where Austria is part of the Bologna Process[11] (Bachelor, Master), Austria is still working on the ongoing process to adjust the system, insofar that they make it comparable to the widely accepted international standard. The importance of particular

---

10 All Explanations according to Austrian Educational System, page 55, provided by Federal Ministry for Education, the Arts and Culture.

Source: http://www.bmukk.gv.at/medienpool/19003/bildungssystem_grafik_e.pdf

11 Bologna Process Website: http://www.ond.vlaanderen.be/hogeronderwijs/bologna/ last online: 28.02.2012.

school types, especially after the 8th grade, will be clearer when looking at the numbers of its visiting students. The structure of the Austrian schooling system can be graphically revisited in the appendix.[12] According to Statistik Austria nearly 1.2 million pupils went to school in 2009/10, excluding universities which would be another 330.000 students. Broken down in Table 1 one can see that the most frequently attended school types after the 8th grade, are academic secondary schools (9 – 12) with a share of 7 %, the vocational schools during apprenticeship (12 %), the three–year type (4,4 %), and the five year type (11.6 %). This paper will deal with two different datasets, which will be explained in detail in the next chapter. For the Programme for International Student Assessment one has to be aware of the different school types that have been tested by the PISA officials. The Programme for International Reading Literacy study which will be the second study to investigate only tested pupils in primary schools. The differentiation between the school types will turn out to be crucial in the analysis.

	Total	Share
Total pupils	1.182.471	
Primary school (Grades 1-4) *	329.440	0.279
Lower secondary school (Grades 5-8) **	217.338	0.184
Special schools[13] (Grades 1-9) **	13.221	0.011
Polytechnics (Grade 9) **	19.315	0.016
New secondary school (Grades 5-8) **	16.848	0.014
Academic secondary school (Grades 5-8) **	114.693	0.097
Academic secondary school (Grades 9-12) **	83.788	0.071
Other types of academic secondary schools **	13.554	0.011
Vocational schools (during apprenticeship) **	140.256	0.119
Vocational schools (3 year type) **	51.712	0.044
Vocational school (5 year type) **	137.534	0.116
Academies	37.354	0.032
Other vocational schools **	7418	0.006

Table 1: Distribution of pupils among school types 2009/2010

a) Data provided by the Statistik Austria
b) Publication: Statistik Austria (2011); Bildung in Zahlen 2009/10 - Schlüsselindikatoren und Analysen.
c) * denotes schools tested in PIRLS; ** denotes schools tested in PISA.

---

12 Austrian Educational System, page 55, provided by Federal Ministry for Education, the Arts and Culture.

Source: http://www.bmukk.gv.at/medienpool/19003/bildungssystem_grafik_e.pdf

13 Specials schools can occur in various forms e.g. inclusive classes for pupils with and without special education needs or classes with support teachers.

4. Dataset Structure

An accurate estimation of an educational production function requires some dataset properties. The criteria, although not easy to fulfill, are relatively straightforward. The most important dataset requirement is to have comparable information on students' performance in schools (e.g. standardized tests). Furthermore, one wants to have data on student and family characteristics such as age, gender and other socioeconomic data. Since, this paper also aims at measuring class size effects, the size of an individual's class should also be given in the dataset. These properties are fulfilled for two large scale projects, namely PISA and PIRLS, whose primary goal is to assess and compare various school system performances of participating countries.

4.1 Programme for International Student Assessment (PISA)

The OECD started to work the PISA in the mid – 1990's and the first survey was conducted in 2000. This large scale project increased its participating countries from 43 in 2000 to 65 in 2009, including all of the 34 OECD countries. In three year intervals pupils at the age 15 and 16 are tested on their mathematic, reading and science skills. The main purpose is to evaluate the knowledge that pupils acquired throughout compulsory schooling. Every three years PISA focuses on one subject in particular, although tests are administered for reading, mathematics and science in every testing year. In 2000 the main focus was on reading, in 2003 it was mathematics and in 2006 it was science. In 2009 PISA's main focus was on reading again. In total 475.460 students have been tested in 2009.

	PISA 2000	PISA 2003	PISA 2006	PISA 2009
Countries	43	41	57	65
Overall Students	228.784	276.165	398.750	475.460
Austrian students	4745	4597	4927	6590
Special Focus	Reading	Mathematics	Science	Reading

Table 2: Overview PISA Studies (2000 – 2009)

a) Data source: see references under OECD PISA Database.
b) Technical reports from http://www.pisa.oecd.org.

The number of students in around 200 Austrian schools has been between 4600 and 5000 in the years from 2000 to 2006, but increased to 6590 in 2009. The data in the table above were calculated directly from the public available PISA datasets at the OECD homepage[14]. In total 7125 Austrian students have been chosen for the PISA study, however, due to cases of illness and other absences 6590 students actually participated. The large increase in the 2009 number of participants can be explained through the fact, that Tyrol and Vorarlberg, which are two federal states in Austria, had a representative study of its population. Hence, more students had been tested than in previous years.

The realization of the PISA test at schools is guided by external trained persons. The selection of students in schools is random. The test lasts two hours where the students have to deal with reading, mathematics and science questions. The questions itself are a mix of multiple choice questions and questions with an open answer format. At the end of test, the students receive a questionnaire on their individual and family characteristics. It is important to note that the PISA tests are in strict confidence. Data are available for scientific intentions, however, the data do neither bare names of pupils, nor school locations. Moreover, if one wants to have data on particular school types (AHS, BMS, HAK, etc.) one has to inquire these data at the Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des Bildungswesens (BIFIE)[15]. For this paper this variable was made available, because it reveals important insights and additionally serves as a robustness check in the analysis. It should be noted, that due to the test design of the PISA study the interpretation of the class size coefficient is not straightforward. The test was designed to evaluate a student's accumulated knowledge up to the testing age. If one had data on the whole schooling career of students including class sizes of previous grades on could estimate a more accurate class size effect. Since students only report the size of their learning group in the subject learning, the effect can only be determined for the grade that they are currently in. Nevertheless, the PISA study serves particularly well to determine the influence of socioeconomic factors and other student characteristics. The second dataset (PIRLS) used in this paper, even though using a different instrumental variable technique already applied earlier by

---

14 http://www.oecd.org/statisticsdata/0,3381,en_2649_35845621_1_119656_1_1_1,00.html. last online: 23.04.2012

15 https://www.bifie.at/ last online: 23.04.2012

Angrist & Lavy (1999), has a clear cut interpretation for the class size coefficient, since they circumvented the endogeneity problem of the class size variable. *"The identification approach exploits the fact that the regressor of interest is partly determined by a known discontinuous function."* (Angrist & Lavy, 1999) Since, class size is not necessarily exogenous to the variation in test scores an instrumental variable estimation is necessary. It might be that though certain school policies, students are selected into particular class within one school. Without an instrument these policies inevitable falsify the results.

4.2 Progress in International Reading Literacy Study (PIRLS)

PIRLS, notwithstanding overshadowed by its big brother the PISA study aims at measuring the reading ability of fourth graders. Based on the Reading Literacy Study of the International Association for the Evaluation of Educational Achievement (IEA) that started in 1970 the PIRLS study was firstly conducted in 2001. The paper pencil test is repeated in five year increments. Recognizing and using acquired information, drawing conclusions, interpreting or linking given information and examining or assessing the content of texts are the main tasks that pupils have to cope with. The PISA study tested randomly selected pupils at the age of 15 or 16 within a school. Regarding the measurement of the class size effect, this is where the advantage of PIRLS comes into play. First, schools were also selected randomly in PIRLS and then all pupils were tested within a particular fourth grade class. Secondly, students as well as parents, teachers and school headmasters received questionnaires to give further background information. Due to the test design it was impossible to collect teacher characteristics in PISA, however it is possible to control for them in the PIRLS dataset. In addition teachers reported the class size, whereas in the PISA study the students reported it. Since, measurement errors concerning class size are less likely in the PIRLS study, the coefficient can be determined more accurately. However, as opposed to the PISA study finding an instrument here is not an easy task. Hence, for instrumental variable estimation the data were aggregated to the class size level and a different approach was applied.

First, the PISA and PIRLS studies are comparable in a several ways. An important feature of both datasets is that they are large scale projects designed to measure individual ability. The PISA 2009 study tested the reading ability of 15 to 16 years old students, whereas PIRLS did the same for fourth graders (i.e. 9 to 10 year old pupils). The OECD mean score for PISA and the International mean score for PIRLS were both 500.. This international mean score deviated slightly over the past years. Although being above the OECD mean, with an average score auf 538, the Austrian pupils were ranked 12th within the 19 participating OECD countries of the PIRLS study.
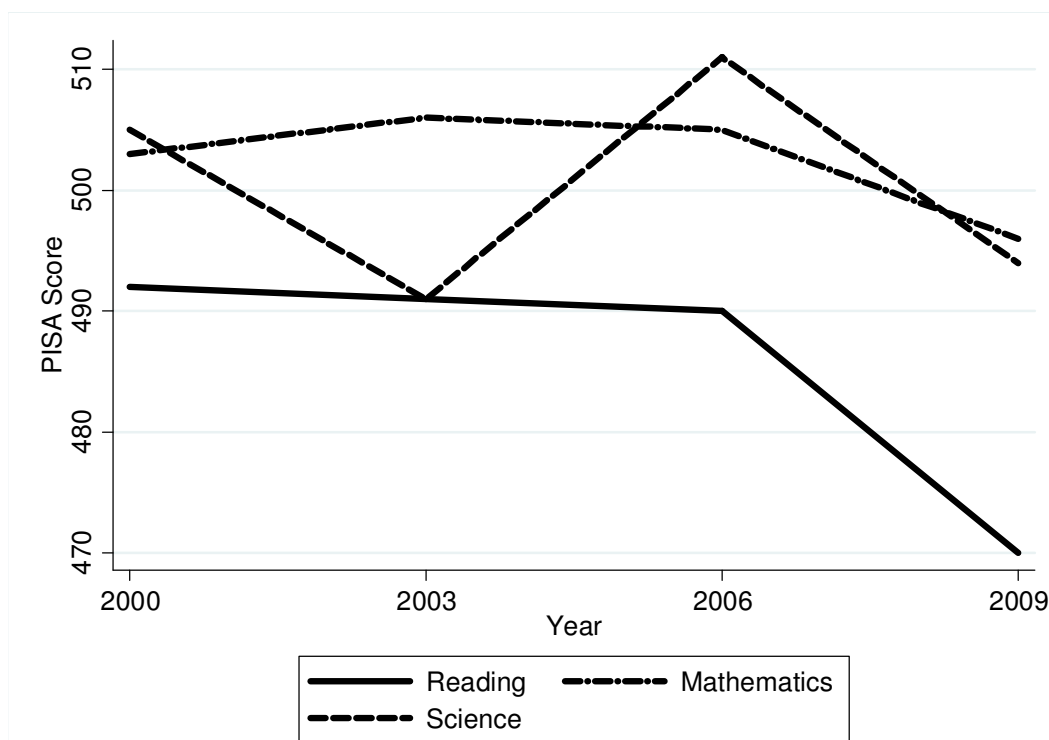
Figure 1: PISA Results Austria for all Subjects (2000 – 2009)

a) Data source: see references under OECD PISA Database
b) All values were weighted by the provided total student weights

In PISA 2003 Austria was ranked 19th out of 29 OECD countries, however, until PISA 2009 Austria dropped to the 31st place of 34 OECD countries. In Figure 1 one can see the results for all three subjects between 2000 and 2009. Comparing the years 2000 to 2009 one can easily see that in all three subjects the scores decreased over time, although the change in science is not significant. The science scores experienced a decline over these 9 years as well, but are far more volatile than the reading and mathematics scores. Especially in reading achievement the Austrian scores decreased.

In 2000, Austria had a reading score of 492, which decreased to 491 and 490 in 2003 and 2006, respectively. In 2009, the scores dropped to 470 points.

Could it be that due to PISA's bad reputation in Austria, the students are not motivated when being evaluated? Baumert & Demmrich (2001) and Brunner et al (2007) found that motivation has no significant impact on scores. Neither, offering money to the students or feedback of teachers, nor being graded by the score on the PISA study altered its outcome. Being part of an International study is motivation enough for the students. PIRLS is conducted every five years and Austria participated in 2006 for the first time.

4.3 Summary Statistics

475460 students in 65 countries were evaluated in PISA 2009. In the summary statistics table 3 data from 6590 Austrian students are reported. The average weighted reading score is 470.28 with a standard deviation of 100.14. The reading score was calculated according to OECD (2009).

"Usually, five plausible values are allocated to each student on each performance scale. Statistical analyses should be performed independently on each of these five plausible values and results should be aggregated to obtain the final estimates of the statistics and their respective standard errors. It is worth noting that these standard errors will consist of sampling uncertainty and test unreliability" (OECD, 2009)[16]. The OECD (2009) also provides an intuition for these plausible values.

"The simplest way to describe plausible values is to say that plausible values are a representation of the range of abilities that a student might reasonably have. (…). Instead of directly estimating a student's ability θ, a probability distribution for a student's θ, is estimated. That is, instead of obtaining a point estimate for θ, (…) a range of possible values for a student's θ, with an associated probability for each of

---

[16]*"The PISA Data Analysis Manual has been developed to provide researchers with various techniques needed to correctly analyse the complex databases. It helps researchers confidently replicate procedures used for the production of the PISA initial reports and thematic reports, and accurately undertake new analyses in areas of special interest. In addition to the inclusion of the necessary techniques, the manual also includes a detailed account of the PISA 2006 database."* (OECD, 2009)

these values is estimated. Plausible values are random draws from this (estimated) distribution for a student's θ". (Wu and Adams, 2002)

Grade average class size in the PISA study is 20.85. This is very close to the Austrian average of 20.83 in 2009[17]. The Austrian average was calculated by collecting data on the total number of Austrian classes and the total number of Austrian students.

PISA 2009	Mean	Min	Max	Std. Dev.	N
Readingscore	470.28	146.53	759.19	100.14	6590
Class size	20.85	1	36	6.25	6190
Female	0.51	0	1	0.50	6590
Grade	9.43	7	11	0.64	6590
German at home	0.89	0	1	0.31	6108
Books at home	3.00*	1	6	1.45	6413
Age	15.81	15.33	16.33	0.29	6590
Father ISEI	44.23	16	90	17.04	5492
Mother ISEI	42.86	16	90	16.00	5305
Student lives w. both parents	0.83	0	1	0.37	6334
Parents education	4.25	0	6	1.24	6276
Home education resources	-0.15	-4.37	0.95	0.96	6457
Home possessions	0.07	-6.82	3.63	0.83	6510
Wealth	0.12	-5.12	2.61	0.82	6503
Grade average class size	20.85	4	34	4.69	6190

Table 3: Summary Statistics PISA 2009

a) Data source: http://pisa2009.acer.edu.au/downloads.php.
b) All data were weighted by a total student weight.
c) The reading score was estimated according to OECD guidelines (OECD, 2009).
d) * means the median is given instead of the mean, because of the variable structure. Std. dev. is still calculated from the mean. The value 3 means that the median student has between 26 and 100 books at home.
e) The variable class size is not actual class size, but approximates the size of a students' learning group in German.
f) N gives the number of observations and is unweighted.

Dividing the total number of students by the total number of classes yields the Austrian grade average class size. In Figure 2, two different types of historic class size trends are reported. The data for these values were all calculated as described above. Class size overall is reporting the average class size for all school types. Class size (primary schools) is a special case of the overall line. As is commonly known, the average educational level, increased strongly in the past decades. Increased demand in higher or more education is also reflected in the class size trend. Whereas in 1923, which is the

---

17 Data downloaded from http://www.statistik.at/web_de/statistiken/bildung_und_kultur/formales_bildungswesen/schulen_schulbesuch/index.html. Since lately only data for the years 2010/11 are available. Hence, these data have to be requested directly at the Statistik Austria Institute.

year where the first class size data are available from Statistik Austria, nearly 75 % of all pupils went to primary school, the same ratio was 28 % in 2009. In 1923 only 136.736 pupils went to some form of secondary schooling. The number of pupils attending schools at the age of 15 and 16, as evaluated by the PISA study, was nearly four times higher in 2009. Overall, average class size went down steadily in the past decades.
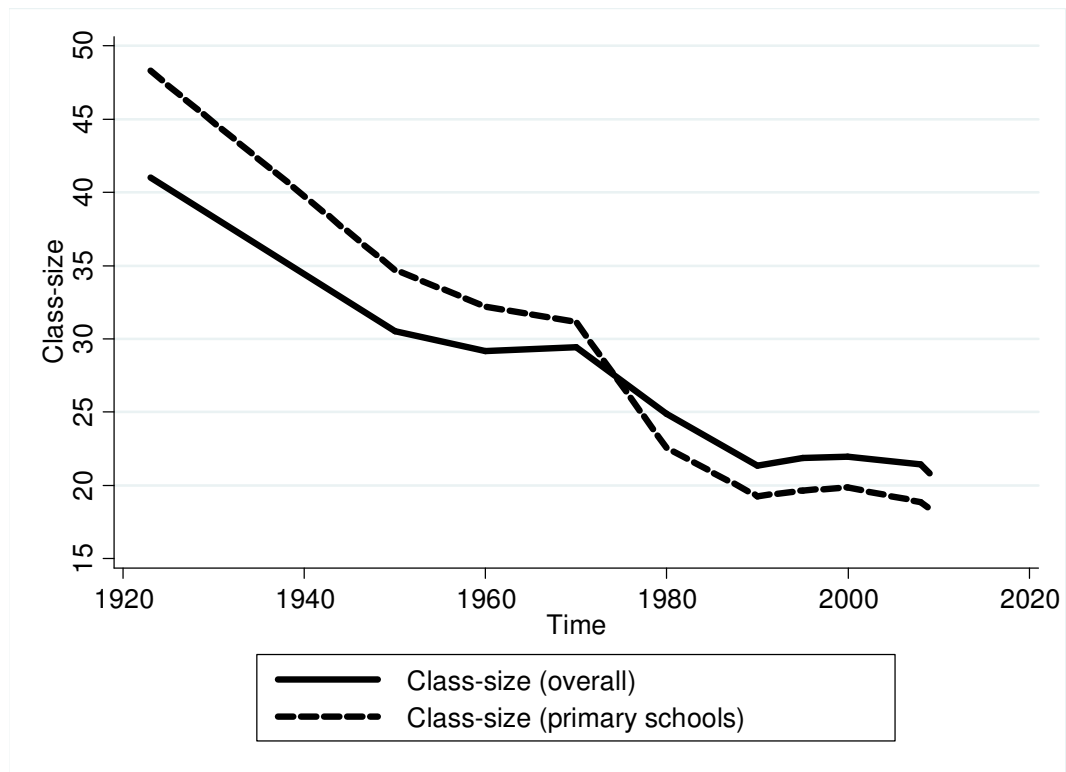

Figure 2: Historical Class Size Trend

a) Data source: Statistik Austria (historical class size data).
http://www.statistik.at/web_de/statistiken/bildung_und_kultur/formales_bildungswesen/schulen_schulbesuch/index.html(23.04.201
b) All values self-calculated.

The PISA class size data reach from 1 pupil (only 1 observation) per class to 36 pupils in a class. Overall, 6190 pupils reported a non-missing value for class size. The share of females in the PISA data is slightly higher compared to males. Since, students were tested according to their age (15 to 16 year olds) the students were also asked to report the grade that they are currently attending. 95 % of the students reported to be in either 9th or 10th grade. Nearly 9 out of 10 of the test subjects reported that their main spoken language at home is German. The students were also asked to report the number of books they have at home. If the *booksathome* variable is equal to 1 this corresponds to

having 0 to 10 books at home and 6 correspond to have more than 500 books at home. The median for the *booksathome* variable is 3.00. This value corresponds to a number between 26 and 100 books. The average student age is 15.81. The difference between the youngest and oldest evaluated students is one year.

Father- and mother ISEI are indicators for the socioeconomic status of the student's family. These indicators range from 16 to 90, where 90 is the highest possible socioeconomic index. These indicators are derived from parents' occupation, education and income. On average fathers have a slightly higher socioeconomic index than mothers. Student lives with both parents is a dummy variable that defines whether a student lives with both parents (1) or not (0). 83 percent of the test subjects reported to live with both parents. The parents' education index ranges from 0 to 6 and is a maximum function of the highest education of father and mother, where the values 0 to 6 represent the International Standard Classification of Education (ISCED).

The three indicators, namely home education resources, home possessions and wealth provide additional measures of students' socioeconomic status. *"The home education resources index includes measure of the existence of a desk, a quiet place to study, a computer that students can use for schoolwork, educational software, books that help with students' school work, technical reference books and a dictionary. The family wealth index includes measures of other goods that might be at a students' home such as a TV, Internet or a dishwasher"* (OECD, 2010). Grade average class size is crucial regarding the instrumental variable measurement framework later in the paper.

If any value in the dataset was not correctly specified for a particular variable, then this value was replaced to missing. The summary statistics give an overview over the whole PISA dataset. Due to misreports not all of the 6590 observations can be used in the regression analysis.

Let us turn to the class size distribution of the PISA data in Figure 3. First, note that the distribution is beginning at 1 and ends at 36, with two larger peaks at a size of 16 and 17. At the center of the distribution there are only half that much classes. Another peak

of the distribution is reached at 24 and 25, following a sharp decrease. The unweighted data in the PISA study are likely to over represent Tyrol and Vorarlberg. Hence, one has to inspect the quantiles of the weighted data. 75 percent of the classes have a class size equal or greater than 20. The median of class size is 23 and 25 percent of the classes have more than 25 students. Ten percent of the classes are larger than 27 students. Comparing the weighted and unweighted data, however, reveals that the difference in class size is of minor importance.
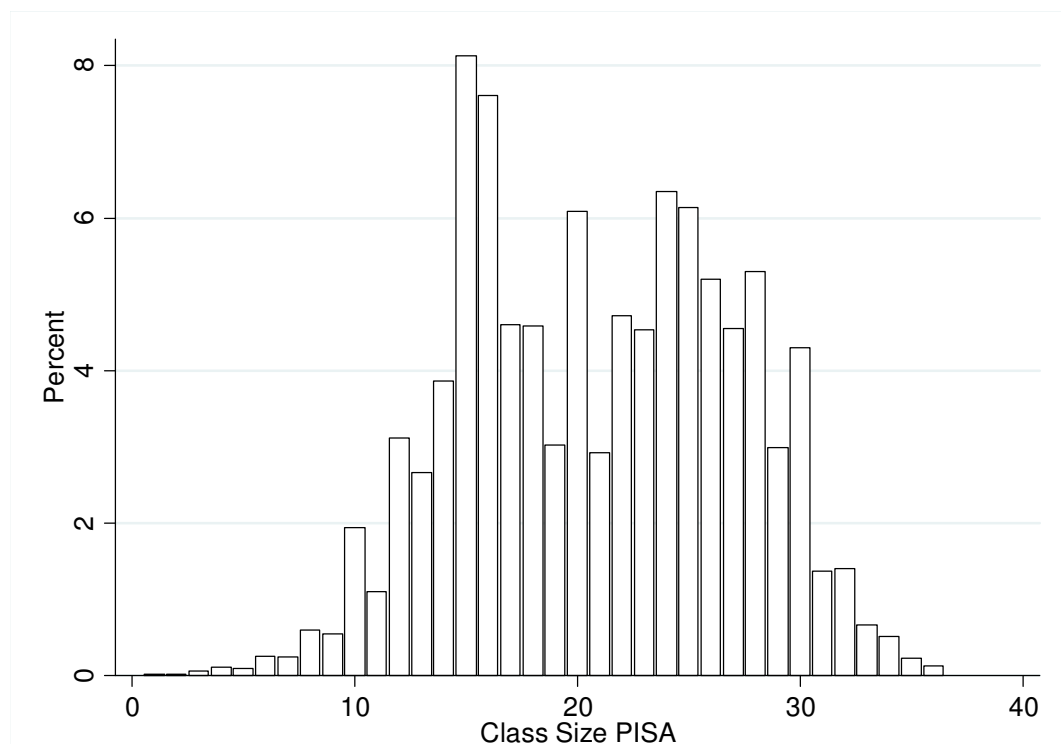
Figure 3: Class Size Distribution PISA 2009

a) Data source: http://pisa2009.acer.edu.au/downloads.php.
b) Class size data is unweighted in histogram.

The legal framework in Austria concerning class size is relatively simple. Depending on the school type a maximum and a minimum class size rule is mandatory. For all school types the minimum class size rule is twenty students in one class, except primary schools where the number of ten students in a class should not be undershot. Until the 8th grade there is a maximum class size rule of 25. After the 8th grade the maximum class size is 30 (§14 SchOG, §21 SchOG, §33 SchOG, §43 SchOG, §51 SchOG, §57 SchOG, §71 SchOG, §100 SchOG, §108 SchOG). One question arises when looking at the

distribution. Why does the class size number 16 and 17 occur that often, if typically, average class size is higher and the legal framework prohibits an undershooting of the lower bound of 20 students in one class?

In the regression analysis this representative sampling of the Tyrol and Vorarlberg should not influence the outcome, since total student weights were applied. Another explanation might be that classes are separated after they exceeded a certain amount of pupils. In Austria this number is 31[18], which might also explain the peaks in the distribution. These weights represent the total student population of Austria and weigh each observation accordingly. The histograms were plotted without using sample weights.

Similar reports for the PIRLS study can be seen in Table 4. The reading score is 538.44 with a standard deviation of 63.63 and the total sample size is 5093.

PIRLS 2006	Mean	Min	Max	Std. Dev.	N
Readingscore	538.44	283.75	725.04	63.63	5093
Class size	21.11	4	30	4.49	4982
Female	0.50	0	1	0.50	5093
Non-native parents	0.17	0	1	0.37	4968
Booksathome	3.00*	1	5	1.22	4802
Childbooksathome	3.00*	1	5	1.16	4803
Age	10.33	9.25	13.16	0.45	5093
Father Education	3.64	1	7	1.41	4578
Mother Education	3.40	1	7	1.19	4659
Highest parental education	3.88	1	7	1.45	4748
Teacher experience	22.19	1	40	10.01	5045
Teacher female	0.88	0	1	0.32	5056
Home education resources	1.95	1	3	0.28	4825
Wealth	2.88	1	5	0.87	4736

Table 4: Summary Statistics PIRLS 2006

a) Data source: http://timssandpirls.bc.edu/pirls2006/user_guide.html.
b) All data were weighted by a total student weight according to PIRLS technical report.
c) The reading score was estimated according to IEA guidelines (Foy and Kennedy, 2008).
d) N gives the number of observations and is unweighted.
e) * means the median is given instead of the mean, because of the variable structure. Std. dev. is still calculated from the mean.

---

18 http://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10009511

Class size is slightly higher in the PIRLS study compared to the PISA study with 21.11, which can be explained by the Austrian legal framework on class size.

50 percent of the tested students are female. 17 percent of the students have immigrant background. The average student age is 10.33 and average teacher experience is 22.19 years. Nearly9 out of 10 teachers are female. The fathers of the students have a mean education of 3.64 and mothers have a mean of 3.40. The wealth as well as home education resources variable serve as a control variable to estimate the others more accurately. The distribution of class size in Figure 4 compared to the PISA distribution is also different in the fourth grade of primary schools. Until the maximum class size of 25 the distribution is increasing steadily. Then there is a sharp drop in the distribution – a fact that will be exploited in the second last section of this paper.
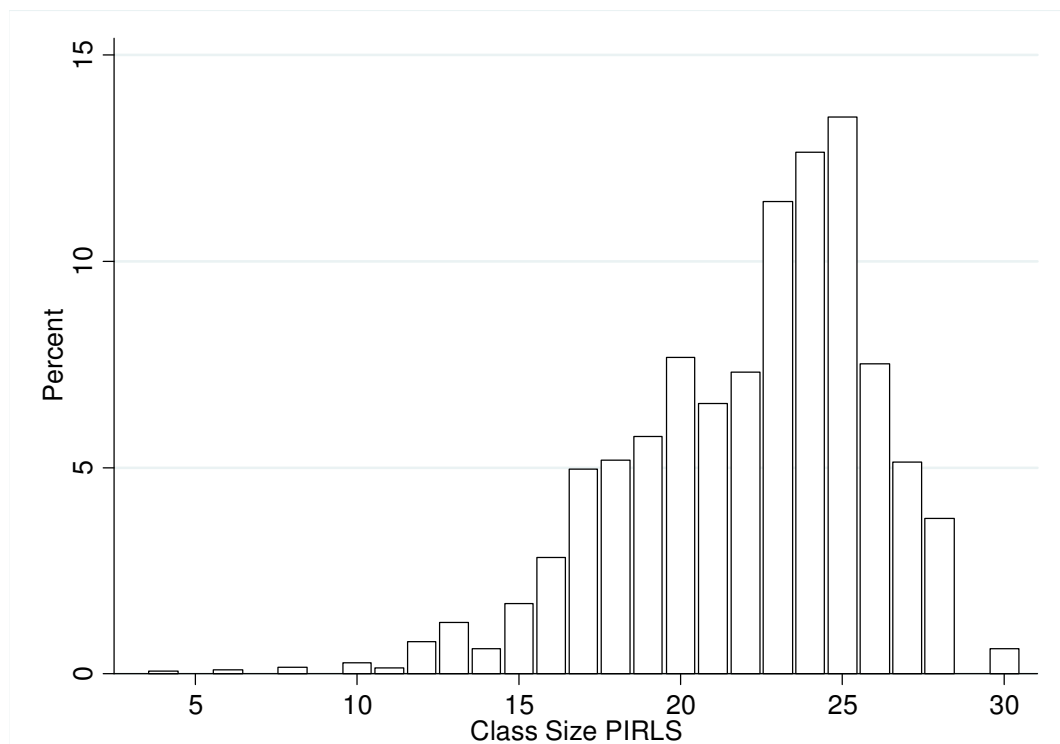

Figure 4: Class Size Distribution PIRLS 2006

a) Data source: http://timssandpirls.bc.edu/pirls2006/user_guide.html (23.04.2012)
b) Class size data is unweighted in histogram.

5. Measurement Framework

Pursuing the ultimate goal of estimating the class size effect in Austrian schools, but also unveiling other determining factors of the educational production function I want to start with a very naïve model and refine it step by step. The starting point is the following model:

$$Y_{icgs} = \alpha + \varphi_{WLS} Size_c + \beta Ctrl_{icgs} + \gamma G_g + \varepsilon r_{icgs} \quad (1)$$

where *Y* represents student *i*'s test score in class *c*, grade *g* and school *s*. The test score will be regressed on a constant *α*, on *Size* which represents the self-reported class size of the students in class *c*, on a control vector *Ctrl* which includes family as well as student characteristics, on a grade variable *G* to control for between grade variation and on some error term. The control vector *Ctrl* includes characteristics such as the gender of the student, measures of socioeconomic status, home possession of students (e.g. books, computers, dishwashers etc.), student's age or parental education. The grade variable will factor out the effect between grades since, typically students in a higher grade perform better on the tests.

Clearly model (1), although often used in the literature is a naïve estimate of the educational production function. A problem that one faces with this naïve estimate is that class size is not necessarily exogenous to the variation in test scores. First, in the rather complex Austrian schooling system there are differences in school types. While an academic secondary school is being considered as a type school which prepares students for college, lower secondary school is more often attended by pupils who are planning to go to work after compulsory schooling. Because of limited entrance it is not unusual that academic secondary schools require students to fulfill certain acceptance criteria.[19] These requirements depend on the schools themselves. Some require students to have at least a "Gut"[20] in German, Mathematics and English – others might even accept students with a "Befriedigend" in certain subjects.

---

19 http://www.bmukk.gv.at/schulen/service/schulinfo/aufnahme_ahs.xml last online: 24.04.2012.

20 Gut corresponds to a "B" in Anglo-Saxon countries; Befriedigend corresponds to a "C".

Evidently, across school ability sorting takes place after the 4th grade. If class size varies between different school types one will not be able to estimate an unbiased coefficient.

It might also be that the parents of good performing students choose to live in areas where average class size is low to boost their children's education even further. On the other hand it might be the case that parents of low performing students choose areas with a low average class size so their kids get extra attention. A priori it is not clear, which effect dominates and whether one would over- or underestimate the class size coefficient. Hence, one has to control for these potential biases by including school dummies or likewise estimate the same model with school fixed effects. Including school dummies estimates a different constant for every school. Whereas without these dummies the model estimates only one constant, by including them a different constant is estimated for every school – 282 in PISA and 160constants in the PIRLS study. Hence, school dummies control for between school variance. Depending on the homogeneity of school types in a system, the between school variance might vary significantly. Due to the broad range of school types in Austria one could anticipate a rather larger between school variance. If student ability selection in particular school types is rather narrow then within school variance, meaning the unexplained differences of students within one school, may be lower.

$$Y_{icgs} = \alpha + \varphi_{SFE} Size_c + \beta Ctrl_{icgs} + \gamma G_g + \delta SD_s + \varepsilon r_{icgs(1)} \quad (2)$$

Including the school dummies δSD_s eliminates all across school variation. The advantage of this strategy is that, no matter whether the school is located in a rural or urban area or whether it is a grammar or lower secondary school– it gives a more precise estimate of the class size coefficient, since we are only comparing classes within a particular grade within one school. Although school fixed effects do not reveal the impact of certain school characteristics on scholastic achievement, they are the most all-encompassing measure of school quality.

Including school fixed effects, however, is not a good strategy when it comes to measuring the impact of socioeconomic factors. Student populations are likely to vary across schools. Student and family characteristics may be correlated to the school

choice itself. Hence, if different school types attract different student populations the effect of individual characteristics will be biased. This means, by comparing models (1) and (2) one could get interesting insights, only by comparing the change of the effects induced by the inclusion of school fixed effects. The school fixed effects model requires data on more than one class per school.

5.1 Within School Sorting Bias Identification Strategy PISA

According to a variable called ABGROUP (see description in the Appendix) in the PISA dataset a common strategy in schools is to assign students according to their ability to certain classes. These sorting effects can occur in two ways. Either a school places worse performing students into smaller classes enabling the teacher to better focus on individual weaknesses, to enable the stragglers to catch up with the better students – or a school places better students into smaller classes to foster their abilities even further.

Thus, class size cannot be considered as exogenous anymore. These school policies can hardly be observed and the bias that comes along with school intern student placements is known as the within school sorting bias. The fact that schools may pursue different goals of student support renders an a priori prediction of the direction of the bias impossible. Hence, the class size coefficient of a school fixed effects model cannot be seen as a lower, or upper bound.

$$\varphi_{WLS} = \gamma + \beta_b + \beta_w \ (3)$$
$$\varphi_{SFE} = \gamma + \beta_w \ (4)$$

In equation (1) the coefficient of class size φ_{WLS} was determined by the actual effect γ, the between school sorting β_b and within school sorting bias β_w, represented in (3). Including the school dummy variables, δSD_s, eliminates any systematic correlation between school variation in student performance leaving behind only the within school sorting bias (4).

Unraveling the within school sorting bias requires a correction of the endogeneity of class size. Endogeneity biases typically call for an instrumental variable. Such an instrument has to fulfill two basic properties.

1) The instrument Z has to be correlated with the endogenous variable φ_{SFE} (i.e. $Cov(Z, \varphi_{SFE}) \neq 0$)

2) The instrument Z must not be correlated with the error term $\varepsilon r_{(1)icgs}$ (i.e. Z must be exogenous or $Cov(Z, \varepsilon r_{(1)icgs}) = 0$)

A common instrument used also by Wößmann & West (2002), Akerhielm (1995), Denny & Oppedisano (2010) is grade average class size within schools. Grade average class size is calculated by averaging over the class sizes within one grade in one school. Using grade average class size as an instrument requires certain dataset properties. For this instrument to work one, first, has to include school fixed effects into the model. In addition to the school fixed effects model requirement of having at least two classes per school one has to have data on more than one grade per school to use between grade variations as a viable source of identification– a property that is fulfilled in the PISA dataset. In a two-step estimation procedure the endogeneity bias can be eliminated. Therefore one has to predict grade average class size in the first stage and use it as a source of identification for the second stage.

The second stage of the two-step (2SLS) estimation procedure with the unbiased instrumental variable (IV) estimator is then given by:

$$Y_{icgs} = \alpha + \varphi_{IV} Size_{cgs} + \beta Ctrl_{icgs} + \gamma G_g + \delta SD_i + \varepsilon r_{icgs(2)} \ (5)$$

where φ_{IV} is an unbiased coefficient for class size and $Size_c$ is predicted by the first stage regression (6), where class size was predicted by grade average class size and the between grade variation is absorbed by grade dummy G_g.

$$Size_{cgs} = \alpha + \phi AvgSize_g + \beta Ctrl_{icgs} + \gamma G_g + \delta SD_s + \varepsilon r_{icgs(3)} \quad (6)$$

Intuitively this strategy asks whether students performed better or worse in the same school in two different grades. If this strategy works, then one can adequately control for the endogeneity bias in the model and remains with an unbiased class size coefficient.

$$\varphi_{IV} = \gamma \quad (7)$$

Why does grade average class size work as an instrument or why are the instrumental variable properties fulfilled?

The first requirement $Cov(Z, \varphi_{SFE}) \neq 0$ is trivial. Since school fixed effects are used, the average is taken over only a handful of classes and as long as class size is distributed over a fairly narrow interval the correlation is expected to be high (i.e. not a weak instrument). This can be verified by looking at the coefficients of the first stage regression. If the coefficient of average class size is non-zero and significant then this holds.

The second requirement, namely, why grade average class size should be exogenous cannot be tested. The argument, why it has to be exogenous, is relatively simple. Calculating grade average class size, by definition, means that the class size is the same in each grade in one school. Including school fixed effects causes a comparison at the school level. Hence, the estimates will be unbiased as long as class size varies across grades (i.e. φ_{IV} is non-zero).The disadvantage of this strategy is that, due to the dataset requirements 23 % of the PISA data observations cannot be used. Schools where only one grade or only one class per grade has been tested have to be dropped from the sample.

5.2 Within School Sorting Bias Identification Strategy PIRLS

In contrast to PISA the grade average class size instrumental variable approach is not applicable in the PIRLS study, since it exploits differences in class size between grades. In PIRLS only one grade was observed, hence there is no between grade variation in class size. A fairly new econometric approach is a regression discontinuity design developed by Twistlewaite & Campbell in 1960. In the economic or rather the econometric literature this approach was introduced at the end of the 1990's. Angrist & Lavy (1999) were the first economists using this approach to measure class size effects in Israeli schools.

They explain it in their paper as follows: *"The approach taken here exploits the fact that the regressor (class size) is partly determined by a known discontinuous function of an observed covariate (school enrollment)"* Angrist & Lavy (1990).

This means that a specific education policy yields exogenous variation in class size, which can be used as a source of identification. Total school enrollment basically determines the size of the class. The authors argue that less populated regions in Israel, where class size is smaller, are inhabited by a greater share of poor families, compared to urban areas which would undoubtedly lead to biased results. Hence, they include a measure of the percentage of students that are disadvantaged in one school to control for socioeconomic differences across the country. It turned out to be crucial in their analysis.

Israeli schools had a maximum class size rule of 40. If there were 41 students in a school two classes had to be created. This rule creating some unpredictable random variation in class size can be exploited to estimate an unbiased class size coefficient. If parents are not exploiting this rule (i.e. are not choosing schools or areas where class size is low) then this strategy is a viable source of identification. Parents' lack of information on school enrollment and additional costs for moving to another area make such a scenario rather unlikely.

An example might help to understand this strategy but first the underlying functions have to be considered:

$$(i) \; e\,/\,[\,floor((e-1)\,/\,40+1)]$$

$$(ii) \; e\,/\,[\,floor((e-1)\,/\,25+1)]$$

In Figure 5 total enrollment in a school is reported on the horizontal axis and class size is reported on the vertical axis. Maimonides rule, which was applied in Israeli schools, is represented by function (i). Up to the first point of discontinuity at 40, class size increases identical to total enrollment. If total enrollment reaches 41 then two classes have to be created, hence class size drops to 20.5. Then for an additional student, average class size increases by 0.5. The same discontinuity arises again at an enrollment size of 81 and 121. In their data of Israeli school's enrollment size was correlated to test scores since socioeconomic status was inversely related to local population density.
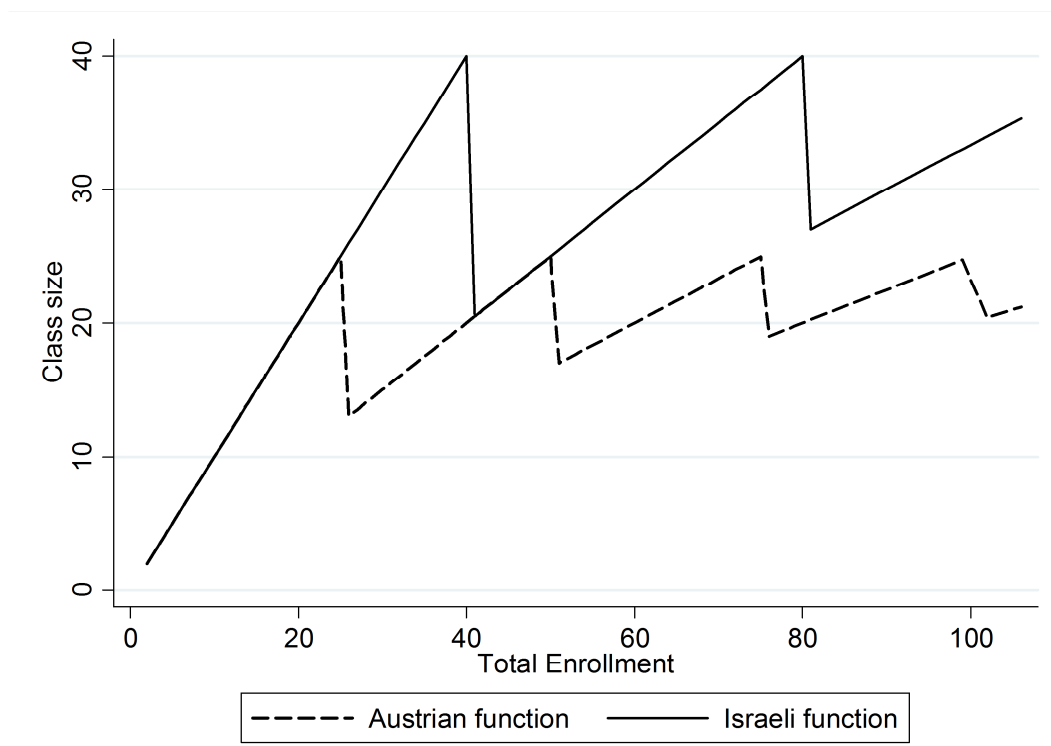

Figure 5: Regression Discontinuity Function

a)   Plots of functions (i) and (ii)

This means that in less populated regions students, where socioeconomic status is low, students performed worse on the tests. The authors argue that *"better schools might face increased demand if parents selectively choose districts on the basis of school quality"* Angrist & Lavy (1990).

The second function in the graph describes the known discontinuous class size function of the Austrian schooling system. The class size cap of 25 is applied in Austrian schools. There might exist schools, which cannot create more classes even though their enrollment size exceeds the limit because of budget constraints. On the contrary some schools might be able to create two classes even under the cap limit because they are privately funded. The model for the estimation on the class level is of the following form:

$$\bar{y}_{sc} = X'_s \beta + n_{sc}\alpha + \delta_s + [\mu_c + \bar{\varepsilon}_{sc}] \quad (8)$$

where $\bar{y}_{sc}$ denotes the average test score of class c in school s, X'_s is a vector of control variables such as enrollment size, n_{sc} determines the class size, δ_s is a random school component that captures the correlation between class averages within schools and $[\mu_c + \bar{\varepsilon}_{sc}]$ is the class level error term. This model forms the basis for naive OLS regression which might not have a causal interpretation; however it also describes the second stage for the IV estimation. The instrumental variable technique under consideration requires the following first stage:

$$n_{sc} = X'_s \pi_0 + f_{sc}\pi_1 + \xi_{sc} \quad (9)$$

where class size is predicted by a known discontinuous function f_{sc} of enrollment size, X' is the same vector of control variables as in (8) and ξ_{sc} is the error term. The vector X' will include measures of total enrollment as well as an index of pupil's socioeconomic status in school s. After predicting average class size by the discontinuous function one can estimate the second stage. It should also be noted that a similar variable to the percent disadvantaged index of the Angrist & Lavy (1990) study will be included.

6. Estimation Results

6.1 PISA Results

Following the measurement framework section I will provide regression results for each model. Although the instrumental variable estimation will provide the most complete measure of the class size coefficient, the foregoing models add useful information on student and family characteristics and also point out the importance of more sophisticated models, since the class size coefficient changes significantly as the measurement strategy improves. The results of models (1) and (2) on the effect of school choice, school types and individual characteristics in Austria taking into account the naïve estimation method is reported in Table 5.

First of all, note that in all the regression equations a total student weight has been taken into account. This total student weight is a crucial application for the Austrian sample to correct for a right distribution of with the school types and the share of females. The estimation procedure was applied as suggested by the OECD. To estimate accurate standard errors replicate weights (W_FSTR1 – W_FSTR80) were used. This procedure ensures that any estimates are unbiased (OECD, 2009; p. 39). Chapter 2 describes the analysis procedure and chapter 3 the two stage sampling design. The replicate weights are then described in chapter 4 (OECD, 2009).

Column 1, representing the simplest model, shows that class size is positively correlated to reading - test scores. First of all note that an increase in class size by 10 would result in an increase in test scores by 20 points. Note that all quantitative interpretations are meant to be ceteris paribus. This means, that the statements only hold under the assumption that all other variable values are held equal. The grade variable is positively correlated with test scores. A student in 10th grade performs 32 points better than a student in 9th grade. Age is not statistically significant, but if anything, more likely to be negatively associated with test scores. Females at the age of 15 and 16 perform 28 points better than their male counterparts when not controlling for school fixed effects. Another interesting finding is the coefficient of the *booksathome* variable with a value of 20.92. Below the coefficient values the t-statistics are reported and according to them the significance levels were determined.

	1) WLS	2) WLS SFE	3) WLS SFE	4) WLS
Dependent Variable	Read. score	Read. score	Read. Score	Read. Score
Class size	2.01**	-0.25	-0.29	1.96**
	7.16	-1.02	-1.13	6.66
Grade	32.50**	34.41**	34.98**	27.43**
	8.49	11.20	11.05	7.12
Female	28.36**	15.72**	16.06**	30.73**
	6.56	5.84	5.82	7.32
German at Home	19.49**	26.18**	26.42**	20.55**
	3.57	6.27	6.22	3.80
Books at home	20.92**	9.40**	9.74**	20.53**
	11.39	7.39	7.57	12.11
Age	-3.47	-1.75	-0.78	-3.04
	-0.49	-0.30	-0.13	-0.42
Student lives with both parents	-3.07	-4.97	-5.21	-3.12
	-0.81	-1.68	-1.71	-0.79
Parental education	9.83**	1.80	1.92	8.63**
	7.87	1.79	1.86	7.06
Home education resources	-2.41	-5.58**	-5.09**	-2.87
	-1.00	-3.19	-2.93	-1.16
Home possession index	21.99**	15.45**	14.61**	21.38**
	3.34	3.54	3.30	3.22
Wealth index	-27.81**	-18.68**	-18.36**	-27.00**
	-5.37	-5.32	-5.18	-5.18
Total School Enrollment				0.025**
				3.14
School Quality				5.03
				1.68
Constant	44.73	4.10	137.54	75.11
	0.48	0.02	0.55	0.79
School Fixed Effects	NO	YES	YES	NO
Observations	5,510	5,510	5,232	5,232
R-squared	0.35	0.63	0.62	0.36
T –statistics reported below coefficients				
** p<0.05, * p<0.1				

Table 5: Weighted Least Square Regression of Models (1) and (2)

As can be seen in the explanation of the variables in the appendix the *booksathome* variable ranges from 1 to 6.The difference between a student who reported to have 0–10 books at home and a student who reported to have more than 500 books at home is on average 104.6 points. Of course the *booksathome* variable is highly correlated to various family characteristics (Van Ours, 2006), but even if one controls for these factors it is an important predictor for students success at school.

For example students whose parents have at most ISCED 4 education reported to have significantly less books at home than students whose parents have better education.[22]Whether a student lives with both parents has no statistical influence on student's performance. Parental education is also significantly positively correlated to test scores. An increase in the variable by 1 (approximately 1 ISCED level) is associated with an increase in test scores by nearly 10 points. *Home education resources* which is highly correlated with the *booksathome* variable has no impact on students success. Lower family wealth as well as home possession is negatively associated with student's success. The wealth variable is a quality measure of home student possessions such as a DVD player or a dishwasher. In general the model, only including student and family characteristics as well as class size, explains about 35 % of the variation in test scores.

Estimating the second model specification, namely including school fixed effects, controls for all across school variation. The school fixed effects model, which is reported in column 2 of Table 5, requires data on at least two classes per school. First note that the R-squared went up to 63 %, meaning that school quality, in fact the most complete measure of school quality, explains 28 % of the variation in test scores in the model. Many might be inferable by comparing the difference in coefficients of column (1) and (2). First it can be discussed to what degree and why the coefficient of class size was altered by the inclusion of school fixed effects. Second, it can be argued, that the student population varies significantly between certain schools. Almost certainly these differences are attributable to differences in school types, rather than differences in schools per se. Hence, the next two subsections will first discuss the change in the class size coefficient and then the change in the other variables between models (1) and (2).

---

22 The mean of the *booksathome* variable is 3.00 if parental education < 5 and 3.65 if education ≥5.

6.1.1 The Change in the Class Size Coefficient

In column 1 the class size coefficient is highly significant and positive with a value of 2.01, but after the inclusion of school fixed effects the coefficient is being reduced to a non-significant value of -0.25. What does account for this difference? As was argued school fixed effects are the most complete school quality measure. Hence, it will be interesting to see how an explicit PISA measure of school quality performs compared to school fixed effects. Furthermore, if better schools have a larger influx of students, which would make total enrollment an implicit measure of school quality, then total enrollment in a school should also be positively correlated with test scores. The third and fourth column of Table 5 report the same school fixed effects regression as before, but with the reduced sample size. In column 3 the sample was reduced intentionally to form a comparable basis for the analysis in column 4. In column 4 school fixed effects were not included, since the coefficient for total school enrollment and school quality would be zero once one controls for all across school variation.

In column 4 total school enrollment and the school quality measure were added to the regression equation and should, as was argued before, implicitly and explicitly control for school quality. Hence, that model should provide comparable results to school fixed effects model if they were a similar complete measure. First, we note that total school enrollment ranging from 15 to 3450 is positively correlated with test scores. An increase in enrollment by one standard deviation of 338.24 increases the average performance of a student by 8.46 points. An improvement of school quality is also associated with an increase in test scores. A standard deviation increase in school of 0.92 increases student performance by 4.63 points.

Comparing the assumed measure of school quality and school fixed effects in terms of the R-squared, emphasizes that the theory of the implicit and explicit quality measures performs poorly. The difference in explanatory power between these two quality measures is 18 % of the total variation in test scores. Additionally the class size variable drops only by 0.05 through the inclusion of the quality measures, but is still positive and highly significant. Hence, it can be inferred that higher quality schools in terms of equipment do not have smaller classes.

In fact one could argue that schools face a trade-off. They could either invest their budget to hire teachers and to build new classrooms, thus reduce average class size or they could invest in school education equipment to increase school quality. Most of the schools will do both to a certain degree, hence the school quality measure, if correlated to class size, should also reduce the class size coefficient significantly, but in fact it does not. Hence, there must be other causes that account for the difference in the class size coefficient. Why did the class size coefficient change that much by the inclusion of school fixed effects?

This drastic change has to be investigated further. Recall the Austrian education system from the section before. In the Austrian schooling system section the various types of schools have been explained. Due to BIFIE policy reasons it is not allowed to associated school types with test scores, since misinterpretations of such results are very common due to complexity of the PISA study. Hence, a different argumentation strategy is necessary.

What we know from before is that AHS have certain entrance criteria. Furthermore BHS are schools where pupils graduate with the A – levels, hence are allowed to go to universities. Suppose the school types in Table 6 below are ordered according to ability of its student population. This is a very strong assumption, however since BHS and AHS have entrance criteria they will only allow some share of the better able students to attend the school. Moreover if these schools attract better teachers as well the level of student performance would spread even further.

Kirabo Jackson (2009) found that in schools with a large portion of minorities teachers are worse. The reason for this is unclear. There might be a direct link between minorities and teachers or it might be the area where these schools are located that teachers want to avoid. Kirabo Jackson (2009) in particular found that high quality teachers tend to leave schools that experience inflows of black students, which is the first link between student population and teacher quality.

Could it be that instead of school quality, the variation in school types accounts for the difference in the class size coefficient and the unexplained variation in test scores in

terms of the R-squared. Therefore the Austrian Institute of Education provided an additional variable, which defines schools by their type. Table 6 shows the mean of class size for all school types and indeed the means diverge significantly.

School Type	APS	BS	BMS	BHS	AHS	Rest
Avg. Class Size	17.99	16.75	21.39	22.62	22.49	13.68
Std. Dev.	5.32	6.19	6.33	6.29	4.45	5.35
Sample Size	809	974	905	1899	1455	148

Table 6: Average Class Size by School Type

a) Data source: Official OECD Database; http://pisa2009.acer.edu.au/downloads.php
b) Data source: School type variable: special data request from www.bifie.at
c) Weighted by total student weights.

The schools that the presumably more able students are attending have larger classes on average. If the assumption holds that students' performance in schools increases from left to right in the table, then the drop in the class size coefficient can be attributed to differences across school types or differences in individual ability which is necessary to get into the school in the first place to some degree. If AHS students perform better on the tests than APS students even though class size is higher, then without the inclusion of school fixed effects one would estimate a biased coefficient.

These numbers point out the importance of the distinction across school types. Depending on the question whether class size has an impact on student's success the question has to be raised why class size is lower in particular types of schools. If class size had an impact on student's success, then lowering class size for less able students can be seen as a redistribution of knowledge or educational fairness. If not, then the question has to be raised why certain school types have a lower class size and infrastructural improvements of school locations should be considered. Under the student ability sorting assumptions the variation of school types must account for some share of the variation in test scores, as well as for the change in the class size coefficient.

6.1.2 The Change in Characteristics

Other interesting findings apart from the class size coefficient can also be inferred from the regression outputs. Starting at the top rows of columns 1 and 2 in Table 5 one has to note that the grade variable as well as the age variable did not change by using school fixed effects. Interestingly the female dummy dropped from 28 to 16. This could also be attributed to the differences in school type argument that was raised in the class size section. It might be that females are less likely to go to schools that prepare students for an apprenticeship. In Table 7 one can see the share of females attending a particular school type according to the weighted PISA data.

School Type	APS	BS	BMS	BHS	AHS	Rest
Share of Females	39 %	39 %	62 %	53 %	59 %	25 %
Sample Size	979	1079	956	1928	1489	159

Table 7: Share of Females by School Type

a) Data source: Official OECD Database; http://pisa2009.acer.edu.au/downloads.php
b) Weighted by total student weights.

Clearly, females are more likely to attend schools that have entrance criteria and less likely to attend schools that prepare for a job after compulsory schooling. Leitner (2001) found that women and men are not only separated into different professions, but rather the range of female jobs is smaller. Whereas 50 percent of Austrian women are concentrated in the largest 4 (out of 27) professions men are concentrated in the largest 7 professions. Moreover except for service and office occupations in typical apprenticeship jobs such as technical jobs or craftman's trade are male dominated (Leitner, 2001). The choice for women after compulsory schooling is limited. Hence, a larger portion stays at school. Moreover men and women seem to choose different types of schools and fields of studies. According to Schneeweis & Zweimüller (2011) these differences in schooling choices between men and women could be reduced, if the share of females in classes is higher. Then, the authors conclude, women would be more likely to choose fields of interest that are typically dominated by males (e.g. physics or mathematics).

It must also be noted that all other coefficients in column 2 and 3 of Table 5 were reduced as well. The argument that school choice does matter might also hold for these

variables. Another crucial question however is, whether it is the school itself that matters or individual ability that selects pupils into these schools. Important to note is that the coefficient of parental education, although still being significant at the 5 % level, dropped by nearly 80 % after controlling for school fixed effects. This could mean two different things. If students self-select into particular school types (i.e. parental background does not influence schooling choice), then it is the school itself that reduces performance differences of different parental educational backgrounds.

School Type	APS	BS	BMS	BHS	AHS	Rest
Avg. Parental Education	3.66	3.87	4.02	4.37	4.92	5.00
Standard Deviation	1.38	1.07	1.11	1.12	1.15	1.00
Sample Size	909	1043	891	1855	1432	146

Table 8: Parental Education by School Type

a) Data source: Official OECD Database; http://pisa2009.acer.edu.au/downloads.php
b) b) Data source: School type variable: special data request from www.bifie.at
c) Weighted by total student weights.

One can easily see that parental education[23] increases from school types that the supposedly less able (always in terms of test scores) pupils attend to the school types that have certain entrance criteria. Whereas average parental education is only 3.87 in vocational schools it is nearly 4.92 in academic secondary schools. It seems as if parental education does play a role, however, not all of the variation can be attributed to it.

Sticking to the R-squared measure of the first two columns in Table 5, one could argue that school quality and total school enrollment could by far not control for the variation in test scores. Hence, for estimating an unbiased class size coefficient one has to apply a school fixed effects model and control for within school variation as explained in the previous section. After the 8th Grade special classes for lower performing students are not common in Austria. Especially in BMS, BHS and AHS such a sorting is very uncommon. These selections usually occur in primary and secondary schools. After talking to teachers as well as students, during research, I found that ability sorting is

---

23 Variable ranges from 0 to 6, where 6 is the highest possible education.

only a common practice in APS and BS. Since APS and BS schools were also tested by PISA officials it is still crucial to control for this possible bias.

6.1.3 Instrumental Variable Estimation Part I

In columns 1 and 2 of Table 9 I applied the instrumental variable technique that is commonly used in the literate, namely predicting actual class size by grade average class size. In column 1 I present the first stage regression, whereas in column 2 the second stage is reported. This instrument should filter out the possibility of within school sorting. In column 1 I regress, as suggested in the methodology section, class size on grade average class size and the predicted values serve as the independent variable in the second stage.

Grade average class size is highly correlated with class size (1.06). Note that below the coefficients the t – statistics are reported, that indicate the significance level of a variable. Hence it can be seen as a strong instrument. The second stage regression reveals that within school sorting does only have a minor effect on the class size coefficient in the PISA study. The decrease in the class size coefficient between the school fixed effects model and the school fixed effects within school sorting bias model is -0.08. The coefficient is still relatively low compared to other socioeconomic variables or student characteristics. It is also insignificant. Compared to column 2 of Table 5 there are only small changes in the other coefficients as well. Note also that the sample is reduced in the instrumental variable approach. This reduction is due to the methodology. Since, the dataset requirement is to have data on more than two classes per school and two grades per school, schools where only one grade was tested had to be excluded from the sample.

Column 3 serves as a robustness check for the instrumental variable approach. In the PISA study schools were asked whether they have some internal policy which suggests an ability sorting of pupils into different classes.

	1) First stage	2) Second stage	3) WLS Rob. Check
Dependent Variable	Class size	Read. Score	Read. Score
Class size			-0.15
			-0.48
Instrumented Class size		-0.338	
		-0.75	
Average Class size	1.06**		
	49.44		
Grade	0.72**	35.57**	31.64**
	4.47	11.52	7.90
Female	0.63**	14.21**	16.94**
	3.98	4.71	4.00
German at Home	0.13	26.60**	29.92**
	0.51	6.00	6.02
Books at Home	0.08	9.61**	10.48**
	0.98	6.67	5.79
Age	0.29	-4.93	1.84
	1.01	-0.81	0.28
Student lives with both parents	0.31	-4.67	-0.40
	1.64	-1.18	-0.09
Parental Education	0.034	0.73	0.27
	0.56	0.61	0.20
Home Education Resources	0.30**	-8.26**	-7.93**
	2.45	-3.85	-3.24
Home Possession Index	-0.725**	16.07**	12.26**
	-2.32	3.02	2.12
Wealth	0.53**	-18.71**	-16.38**
	2.15	-4.01	-3.18
Constant	-13.25**	-20.64	-29.51
	-3.45	-0.08	-0.27
School Fixed Effects	YES	YES	YES
Observations	4,037	4,037	3215
R-squared	0.54	0.56	0.54
T – statistics below coefficients			
** p<0.05, * p<0.1			

Table 9: Instrumental Variable Estimation

Re-estimating the model in column 1 and excluding the schools that reported to have such a school policy yields the coefficients in column 3. There is only a small change in the class size coefficient, however, it is still not significant at the 10 % level. Neglecting the significance level and the assumption of an unbiased coefficient leads to the conclusion that a decrease in average class size by ten pupils increases student performance in the PISA study by a bit over 3 points – a relatively small change. The data also reveal that within school sorting is a minor problem in BMS, BHS and AHS.

School Type	APS	BS	BMS	BHS	AHS
Some within school sorting policy	877	927	229	396	197
No within school sorting policy	56	104	727	1500	1162
Misreports	102	48	0	32	130
Total sample	979	1079	956	1928	1489

Table 10: School Sorting Policies by School Type

a) Data source: Official OECD Database; http://pisa2009.acer.edu.au/downloads.php
b) Data source: School type variable: special data request from www.bifie.at
c) ABGROUP variable serves as a basis.

Whereas in APS and BS within school sorting seems to be a common practice, this can be negated for the other school types. Hence, another robustness check for the class size coefficient could be to estimate a coefficient for BHS and AHS, because it is known that they do not have such a policy. And also for APS and BS this could be interesting when only taking these schools into account who reported not having such a policy. Hence, coefficients were estimated separately for each school type according to the strategy in column 3 of Table 9.

School Type	APS	BS	BMS	BHS	AHS
Class size coefficient	11.24	0.80	-0.39	0.12	-0.18
T– statistics	2.20	0.34	-0.89	0.23	0.21
Observations	41	91	592	1344	1049

Table 11: Class Size Coefficient by School Type (robustness check estimation)

a) Data source that underlies the calculations: Official OECD Database; http://pisa2009.acer.edu.au/downloads.php
b) Data source: School type variable: special data request from www.bifie.at
c) Weighted by total student weights.
d) Estimation procedure according to OECD guidelines.(OECD, 2009)
e) The model also applies the ABGROUP identification of a between class sorting policy.

Class size does not have any effect in either type of school if measured separately. The effect in APS is positive, however due to the very small sample size not meaningful. At the end of this section it is important to consider that the PISA study quite the contrary to the Tennessee Star experiment was not created to measure the effect of class size on student achievement. Its main purpose is to measure the level of knowledge of a student acquired over the total school career. Two interpretations are possible. It might on the one hand be that there is just no effect of class size in 9th and 10th grade. On the other hand it might be that there is no effect in the PISA study, since the study itself is inappropriate for such an analysis. Overall family and student characteristics, as well as school type choices have a strong impact on students' success.

6.2 PIRLS Results

The second study that will be investigated is the Progress in International Reading Literacy Study. Following again the same methodological framework for model (1) and (2) I will estimate a class size coefficient for 4th graders in primary school. The strategy is the same as in the PISA study before. The instrumental variable approach will deviate from the one used before since in PIRLS only 4th graders have been observed and the grade average class size approach requires a more complicated dataset structure, which is not given in PIRLS. Some advantages of the PIRLS study are, that data on teacher characteristics are available and that only one school type was tested, namely primary schools. Therefore no distinction between different types of schools is necessary, which facilitates the determination of the class size effect. Furthermore, the data on class size are more accurate since the teachers reported the class size; on the contrary to the PISA study where the students self-reported it.

In Table 12 the first column corresponds to model (1). First, note that the number of observations in column 1 is 4285 and the R-squared is 0.21. According to this model class size does not have any effect on the readings scores (-0.13). Females do again perform significantly better than their male counterparts, however the difference compared to the PISA study is much lower (28 points vs. 6 points). Age is negatively correlated to test scores, at least in the range 15.33 – 16.33 in the PISA study.

	1) WLS	2)WLS SFE	3)WLS Rob. check
Dependent Variable	Reading score	Reading score	Reading score
Class size	-0.13	-1.72	-0.26
	0.33	-1.21	-0.60
Female	6.20**	3.34	3.36
	2.64	1.38	1.43
Age	-12.75**	-13.67**	-13.56**
	-4.72	-5.49	-6.19
Books at Home	7.76**	7.15**	7.32**
	6.70	6.06	6.21
Childbooks at home	10.03**	8.84**	10.10**
	8.58	7.62	8.75
Homeeducation resources	--18.79**	-16.87**	-17.83**
	-4.07	-3.35	-3.65
Non-native parents	-15.48**	-11.80**	-17.22**
	-4.40	-3.01	-4.26
Highest parental education	2.93**	2.76**	3.47**
	2.97	2.35	3.01
Wealth	-0.07	0.80	0.33
	-0.06	0.65	0.26
Teacher Experience	-0.07	-0.44*	-0.16
	-0.49	-2.14	-0.99
Teacher Female	-1.49	-5.48	--1.96
	-0.27	-0.72	-0.43
Constant	646.75**	689.66**	658.98**
	21.09	14.68	22.33
School Fixed Effects	NO	YES	NO
Observations	4285	3602	3602
R-squared	0.21	0.30	0.23
T – statistics below coefficients			
** p<0.05, * p<0.1			

Table 12: Weighted Least Squares Estimation PIRLS 2006

a) Data source: Official IEA PIRLS Online Database; http://timssandpirls.bc.edu/pirls2006/user_guide.html
b) Weighted by total student weights.
c) Estimation procedure according to IEA guidelines. (Foy and Kennedy, 2008)

A difference in students' age of one year is estimated to lower the test score by 12.75 points. An interpretation could be that, if a student is older than his class members he might be less able in the first place. The *booksathome* and *childbooksathome* variable do have again a large impact on scholastic achievement. The home education resources coefficient which is correlated to the number of *booksathome* increases the effect of educational material at home even further. A student who has non-native parents performs 15.48 points worse than a student with Austrian parents. The more educated parents are the better the students performed on the PIRLS study.[24] The difference between a student where one parental unit has a university degree and a student where one parental unit has at most finished primary school is 11.72 points (4 x 2.93) in the PIRLS study.

Neither the experience of a teacher, or the gender of a teacher, nor family wealth does have any significant influence on test scores. In the second column school fixed effects were included. The number of observation dropped to 3602, because schools where only one class has been tested had to be excluded from the dataset. The model explains 30 % of the total variation of test scores.

Even though the class size coefficient decreased to -1.72 it is not significant at the 10 % level. A decrease, disregarding statistical significance, in class size by 10 pupils would result in an increase of 17 points at the PIRLS study. A relatively small and insignificant change in the coefficients suggests that a reduction in class size, according to these data, is not advisable. Interestingly teacher experience is now negatively associated with test scores. Most likely, this is not due to teacher experience itself, but rather due to teachers' age. This demonstrates that on average younger teachers perform better than their older colleagues. The other coefficients remained at a relatively comparable level to column 1.

The question also remains whether the change in the class size coefficient between models (1) and (2) is due to the inclusion of school fixed effects or due to the necessary reduction in the sample size. Hence, a re-estimation of column (2) is necessary; excluding school fixed effects. It can be seen once one does not control for any across

---

24 Variable ranges from 1 (at least one parent has a University degree) to 5 (one parent has at most primary education).

school variation (e.g. excluding the school dummies) the new class size coefficient is very similar to column 1.

Recalling the changes of the other coefficients in column 2 (e.g. student and family characteristics) between model (1) and (2) in the PISA study we noted that the changes were surprisingly large when controlling for school fixed effects. The story is different for the 4th grade in primary schools. The changes are relatively small in absolute numbers. This also supports the theory that the variation in school types was crucial for the differences in coefficients. An interesting fact in the PIRLS study is, once one controls for school fixed effects females do not perform better than males. In addition there are no significant differences between rural, suburban and urban primary schools once one controls for all the other variables in the model.[25] Whereas in the PISA study females performed significantly better than males in the reading section, in PIRLS females, if anything, only performed slightly better.

Lynn and Mikk (2009) found that one explanation of the increase in reading ability gap is due to a deeper engagement in language related abilities for females. By comparing summary statistics from the PIRLS studies for all the participating countries they found, that even though boys on average do own more books, boys and girls read different kinds of books. Girls read more poetry, popular fiction and romance books and boys are more likely to read articles like sports pages, cartoons, comics, news and science fiction. The main question is why this gap is widening over the years?

The Canadian council on learning (2009) found that girls tend to do more non-assigned readings and they are more likely to read for enjoyment. Boys, however, have different hobbies. They are more likely to watch television or a movie, so they continue. Since skill comes with practice and women read more than men the difference will widen over the years. Thus, the Canadian council on learning (2009) suggests an altering of boys' attitude towards reading at home, as well as in school. *"Parents should encourage their children to read more books at home instead of watching television* (Canadian council on learning, 2009)." Although, changing the reading attitude is not an easy task, a starting point could be to provide books that boys enjoy to read.

---

25 School area variables (rural, suburban, urban) are not reported in the tables. Calculation have been performed, however none of these values was statistically significant and relevant for the results.

6.2.1 Instrumental Variable Estimation Part II

In the last section I estimated a class size coefficient for fourth graders, taking into account student, family and teacher characteristics as well as across school variation. The remaining step to estimate an unbiased class size coefficient is to control for within school sorting in primary schools. In the measurement framework the identification strategy has been discussed. As was explained previously, in Austrian primary schools the class size cap in primary schools is 25. If there are more than 25 pupils in one class, an additional class has to be created.
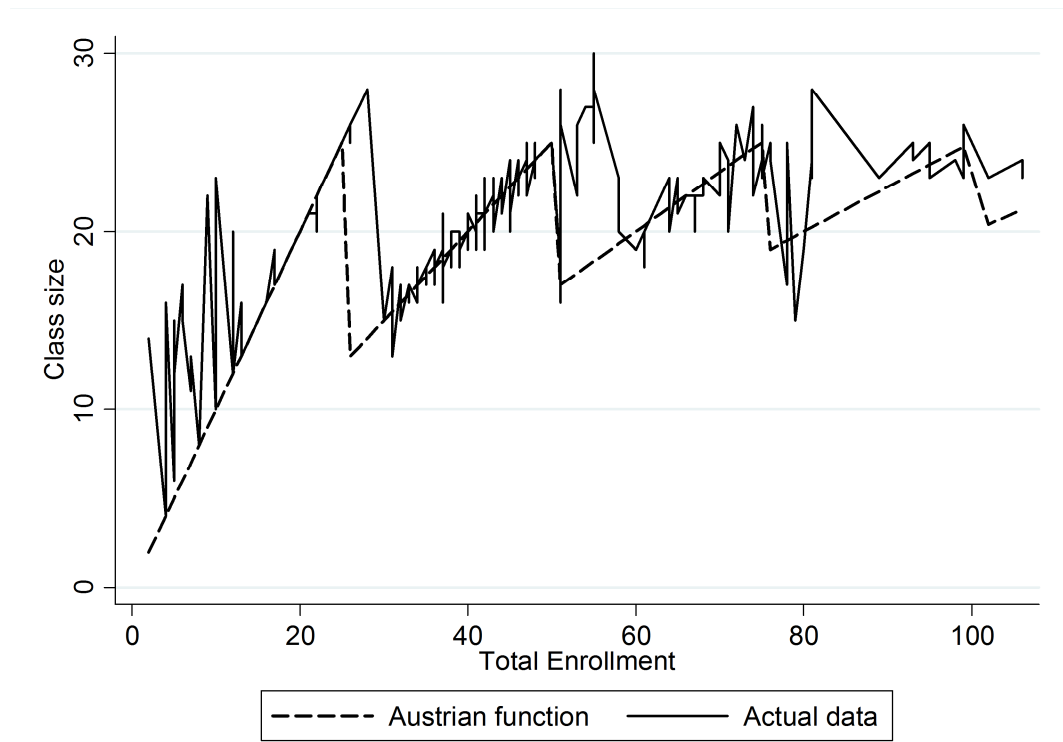

Figure 6: Regression Discontinuity Function and Actual Data

a) Data source: Official IEA PIRLS Online Database; http://timssandpirls.bc.edu/pirls2006/user_guide.html (23.04.2012)
b) Actual data were calculated directly from the dataset by aggregating individual class size in each class

On average the cap rule of 25 is a good approximation as you can see in Figure 6, where the Austrian class size function and actual aggregated data were plotted. It can be seen that the actual data are well explained by the discontinuous function. A numerical relation of these plots will be estimated in the first stage regression.

Because of the necessary aggregation to the class size level the number of observations drops to 243 (i.e. the number of classes tested). The data were aggregated according to an identification variable that indicated which students are in the same class. The regression output which is considered first is a simple WLS estimate. In column 1 the simple WLS estimate indicates that class size is negative with a value of -0.81, but has no statistical significant effect on student outcome. The percent disadvantaged index

	1) WLS Estimate	2) First stage IV	3) Second Stage IV
Dependent Variable	Avg. Score	Avg. Class size	Avg. Score
Function		0.40**	
		6.52	
Instrumented class size			-4.34*
			-1.73
Avg. Class size	-0.81		
	0.89		
PD	-4.97	0.86**	--6.58*
	-1.37	2.83	-1.82
Total Enrollment	0.06	0.05**	0.48*
	0.52	3.35	1.90
Constant	560.49**	9.53**	619.65**
	32.99	11.04	15.92
Observations	243	243	243
R-squared	0.04	0.51	0.09
Robust standard errors below coefficients			
** $p<0.05$, * $p<0.1$			

Table 13: Regression Discontinuity Instrumental Variable Estimation

a) Data source: Official IEA PIRLS Online Database; http://timssandpirls.bc.edu/pirls2006/user_guide.html

ranging from 1 to 4^{26} suggests that students in a school where less or equal than 10 percent of its population are disadvantaged on average perform 5 points better than in a school where between 11 and 25 percent are disadvantaged. Furthermore, the simple model indicates that there is no relation between total school enrollment and performance on tests. The regression, however, explains only 4 % of the total variation in aggregated test scores. In column 2 the first stage of the regression discontinuity instrumental variable approach is reported. The known discontinuous function is correlated to grade average class size. The correlation fulfilling and proving the first

---

26 Economic disadvantage index: 1= 0-10 %; 2=11-25 %; 3=26-50 %; 4 = 50 % or more.

requirement of a valid instrumental variable, with 0.40, is also far from being a weak instrument and highly significant.

In the third column, namely the second stage regression, the effect of the instrument compared to first column nearly decupled to -4.34. The effect is significant at the 10 % level and more than doubled compared to the analysis without taking within school sorting into account. The R-squared went up to 9 %. Total enrollment is positively correlated to test scores and the percent disadvantage index went up slightly to -6.58, however is still highly significant. Even though the approaches used in the PIRLS study differ and the R-squared is significantly lower, it can be seen that class size has a positive impact on scholastic achievement. The regression discontinuity design is not possible in the PISA dataset, since individuals were not assigned to particular classes.

Although an instrumental variable technique that was applicable on the individual level would be preferable to the aggregated instrument the regression discontinuity approach yields an unbiased coefficient on the class level. In general it is unclear what happens to the non-aggregated class size coefficient, when one controls for within school sorting. It is likely, since relatively small integrated classes are common practice in Austria that the coefficient would have been lower than before, but this paper could not provide evidence for an effect on the individual level. If the class size coefficient is decreasing further when controlling for within school sorting it is unclear whether it is bigger or smaller than the IV – estimate in column 3 of the regression discontinuity design. This coefficient should however be a good approximation to the unbiased individual coefficient. Hence, the coefficient estimated in the school fixed effects model can be seen as an upper bound for the class size effect.

In the IV estimate a class size reduction of 10 pupils would increase student performance by more than 40 points; ceteris paribus of course. The important remaining question is what the costs of such an intervention are? And could the money be invested in a better way to improve school quality? If classes were reduced by 50 % one would need twice that many class rooms, teachers and also more administrative staff. A cost benefit analysis might shed light on whether a reduction of class size is efficient. This analysis is beyond the scope of this paper and left for future research.

7. Conclusion

Due to difference in schooling systems around the world this papers goal was to determine the effects that influence scholastic achievement of individuals, with a particular focus on the effect of class size. Complex structures of the Austrian schooling system, however, make it hard to take a closer look at the effects of class size on scholastic achievement. Hence, the paper used two different datasets to examine the Austrian educational production function with a particular focus on class size. The PISA dataset was only suited to a limited extent. No effects have been found in terms of class size.

Therefore, two reasons might be possible. Either there are simply no effects of class size on scholastic achievement; or the more likely explanation is that the PISA study is not applicable for this analysis since it the original purpose of this study is to measure the level of knowledge that students acquired up the testing age.

A more accurate class size coefficient could be estimated if one had class size data on the whole schooling career. Moreover students at the age of 15 and 16 are expected to upgrade education on their own. Reading skills are developed way earlier, namely in primary schools and even more important, at home. Clearly the analysis showed that parental education strongly influences the reading ability. Moreover, home education resources such as books or child books are of high value for scholastic achievement. The original source of this effect is unclear, but it is likely that good parental background has a strong impact.

Some evidence for two-tier education society has been found. Since large changes in the coefficients of *education*, *female*, *wealth* and *booksathome* occurred when school fixed effects were included, the question of equality of opportunity has to be raised. The mobility of education seems to be low in Austria (Fessler & Schneebaum, 2012). If a student's parents are not well educated the student is also likely to be not well educated either. The key task for a schooling system is to guarantee equal opportunity for all children. Therefore further evidence on the mobility of education would be necessary. The grade average class size instrumental variable approach was also validated by the robustness check who found no effects of class size on student achievement.

In the analysis of the PIRLS study a different instrumental variable technique was introduced. In fact smaller classes do yield benefits for students. Even though the sample size in the IV – approach is small, the evidence is convincing. Class size does matter. Moreover, although the general belief that females do read better is true, according to our findings this is not true in our fourth grade sample from PIRLS. Practice makes perfect. Boys and girls have different attitudes towards leisure activities (Canadian Council on Learning, 2009). Hence, the ability improves over time and leads to a28 point difference in the PISA study. This is clear evidence that if parents want to foster their kids' reading ability they should provide the books.

Evidence should be collected on the impact of class size in different grades, especially in primary school. It might be that in the very early stage, namely first and second grade, class size is even more important. The foundation of education is constructed early on in life. Hence, the suggestion is to introduce a schooling system with different class sizes at different grades. This might help students to build up a good foundation of basic skills, such as reading or mathematics. Later on in their schooling career students could then start to learn of their own accord if the basics were mediated sufficiently well. If a student has always been a bad reader he will most likely never enjoy reading and never practice it on a regular basis. Hence, it is important to equip the students with the basic tools; i.e. good reading skills.

At the end it should be noted that for both datasets checks on non-linearity (e.g. logarithmic or exponential shape of class size function) have been conducted; no further insights were gained. The explanatory power of the models was weakened when including exponential or logarithmic class size variables. Nevertheless, due to the finding that class size does matter in primary schools further investigation will be necessary. Hence, future work in this area might consider a cost benefit analysis of the effect of smaller classes and the benefits of a stronger differentiation of class size between grades.

8. Zusammenfassung

Aufgrund der Unterschiede zwischen Schulsystemen ist das Ziel dieser Arbeit die für den Schulerfolg ausschlaggebenden Faktoren zu quantifizieren. Ein spezieller Fokus dabei wurde auf die Bestimmung des Effektes der Klassengröße gelegt, welcher aufgrund komplexer Strukturen des österreichischen Schulsystems einer anspruchsvollen Analyse bedarf. Dafür wurden der PISA 2009 und der PIRLS 2006 Datensatz verwendet. Da der PIRLS Datensatz unpassend für die Anwendung der Durchschnittsklassengröße als Instrumentalvariable ist, musste ein anderer Ansatz, nämlich der einer Regressionsdiskontinuität, gewählt werden.

Bei der PISA Studie zeigt sich, dass die Klassengröße keinen Einfluss auf die Schulleistung hat. Dafür kann es zwei Gründe geben. Einerseits könnte die Studie selbst unpassend für eine Analyse der Auswirkungen der Klassengröße auf den Schulerfolg sein, da die PISA Studie die während der gesamten Schullaufbahn erworbenen Kenntnisse der Schüler testet und sich nicht auf das Gelernte in einer bestimmten Schulstufe beschränkt. Andererseits wäre es auch möglich, dass Klassengröße und Schulerfolg in höheren Schulen ganz einfach nicht korreliert sind.

Die PISA Studie ist allerdings sehr aufschlussreich wenn es um die Quantifizierung der sozioökonomischen Aspekte geht. Hier zeigt sich, dass vor allem der Bildungsgrad der Eltern mit allen zugehörigen Aspekten einen starken Einfluss auf den Schulerfolg hat. Auch die verschiedenen Schultypen in Österreich tragen einen großen Teil zur Divergenz in Testergebnissen unter Schülern bei. Dieser Effekt wird sicherlich durch einen Sortierungsmechanimus nach der Volksschule vergrößert. Einerseits wird in vermeintlich „besseren" Schulen ein anderer Lehrplan unterrichtet, aber andererseits ziehen zum Beispiel Gymnasien von Vornherein eine andere Schülerschicht an. Das ist auf die Notenvoraussetzung in der Volksschule und auf die Bildungsschicht der Eltern zurückzuführen, wobei beide dieser Faktoren ebenfalls korreliert sind. Es scheint als wäre die Mobilität von Bildung eher gering in Österreich. Wenn die Eltern gebildet sind dann sind auch deren Kinder gebildet und vice versa.

Bei der Analyse des PIRLS Datensatzes zeigt sich ein etwas anderes Bild. Zwar sind Bildungsgrad der Eltern und andere sozioökonomische Faktoren immer noch bedeutend

in der Bestimmung des individuellen Schulerfolgs, jedoch zeigt sich, dass auch die Klassengröße einen bedeutenden Einfluss auf den Schulerfolg hat. Während bei der Analyse auf Schülerebene kein signifikanter Effekt bei unterschiedlicher Klassengröße festzustellen ist, so wird bei der Anwendung einer Diskontinuitätsfunktion als Instrumentalvariable klar, dass ein positiver Effekt von kleineren Klassen nicht auszuschließen ist. Aufgrund der Aggregierung auf Klassenebene konnte der genaue Effekt auf Schülerebene nicht bestimmt werden.

Es wäre durchaus denkbar, dass kleinere Klassen zumindest in Volksschulen eine Leistungssteigerung bewirken auch wenn die Ergebnisse nicht ganz eindeutig sind. Wenn sich so ein Trend aufgrund weiterer Untersuchungen bestätigen würde läge der Vorschlag einer Differenzierung der Klassengröße nach Schulstufen nahe. Beginnend bei sehr kleinen Klassen in der Volksschule könnte die Klassengröße stetig vergrößert werden. Dies würde einerseits das Bildungsbudget nicht zu sehr beanspruchen und anderseits den Schülern einen besseren Bildungsgrundstock schon früh in der Schulkarriereermöglichen. Somit könnte man etwaige sozioökonomische Nachteile schon früh in der Schulkarriere durch individuellere Betreuung ausgleichen. In höheren Schulstufen kann man laut diesen Ergebnissen, durchaus, größere Klassen einführen. Dafür sollte mehr Wert auf außerschulische Betreuung gelegt werden.

Zum Beispiel könnte man die Schüler auch in Ihrer Freizeit zum Lesen von Büchern anregen. Wenn Schüler schon früh richtig lesen lernen, dann verlieren diese auch vielleicht nicht das Interesse sich ab und an ein Buch zur Hand zu nehmen. Interessant für zukünftige Arbeiten in diesem Bereich, wäre es einerseits die genauen Kosten von kleineren Klassen zu ermitteln. Es ist denkbar, wenn Schüler schon in der Volksschule besser betreut werden, dass diese ein größeres Interesse an Bildung entwickeln und somit auch in höheren Schulstufen einen größeren Wert auf Selbststudium legen.

9. Appendix

9.1 PISA 2009 Variables Description

Note that all misreports in the variables were transformed into missing values and were not included in any of the analysis or summary statistics. The explanations are of the following form: variable name in the paper – variable name in the PISA 09 Dataset.

Readingscore - W_FSTR1 - W_FSTR80: scaled to an OECD-mean of 500 and a standard deviation of 100. It was calculated by using the replicate weights (W_FSTR1 - W_FSTR80) provided in the PISA 2009 dataset. The reading score variable itself is not part of the regression analysis. The analysis is always based on the replicate weights.

Grade - ST01Q01: information about the current grade which pupils attend; ranges from 7 to 11.

Class size - ST35Q01: pupils were asked "On average, about how many students attend your German class?" (class size ranges from 1 to 36)

Female - ST04Q01: 1 if student is female, 0 if male.

German at home - ST19Q01: 1 for those who reported speaking German at home and 0 otherwise.

Booksathome – ST22Q01: Students reported on a scale ranging from 1 to 6 whether they have 0 – 10, 11 – 25, 26 – 100, 101 – 200, 201 – 500 or more than 500 books at home.

Age - AGE: represents student's age (ranges from 15.33 – 16.33).

Father ISEI - BFMJ: Fathers socioeconomic status; this index is derived from fathers occupation and also depends on fathers education and his income; (ranges from 16 to 90) 90 represents the highest socioeconomic status.

Mother ISEI - BMMJ: Mothers socioeconomic status; same as Father ISEI.

Wealth - WEALTH: Index of family wealth; based on responses on whether students have "a room of their own", "a link to the internet", "a dishwasher", "a DVD-player" and three other country specific items, and responses on the number of cell phones, televisions, computers, cars and the rooms with a bath or shower at home.

Student lives with both parents - FAMSTRUC: 1 for those students who live with both parents and 0 otherwise.

Father education - FISCED: Education level of father (ISCED); ranging from 0 to 6, where 0 represents no education, 1 = ISCED 1, 2 = ISCED 2, 3=ISCED 3B or C, 4=ISCED 3A or 4, 5=ISCED 5B and 6=ISCED 5A or 6.

Mother education - MISCED: Education level of mother (ISCED); same as father education.

Highest parental education - HISCED: max (Father education, Mother education)

Home education resources - HEDRES: Index of the educational resources at student's home.

Home possession index - HOMPOS: Students resources such as TV, Internet etc. at home.

Grade average class size: calculated by averaging over students in one school, that reported the same class size.

ABGROUP: provides a measure of whether a school has an internal ability sorting policy; is used for the robustness check estimate.

Total school enrollment - SCHSIZE: Total number of students enrolled in the school where the pupils have been tested.

School quality - SCMATEDU: Index of school quality.

Student weight – W_FSTUWT: Final student weight; the sum of the weights constitutes an estimate of the size of the target population.

9.2   PIRLS 2006 Variables Description

Readingscore: calculated according to PIRLS Data Manual using variables
ASRREA01 – ASRREA05.

Class size - ATBGCSTD: teachers reported the class size; Ranges from 4 to 30.

Female - ITSEX: 1 if student is female, 0 if male.

Age - ASDAGE: represents student's age (ranges from 9.25 to 13.16).

Booksathome - ASBHBOOK: Index of amount of books at home ranging from 1 to 5; 1=0 – 10 books; 5=more than 200 books.

Childbooksathome - ASBHCHBK: Index of amount of childbooks at home ranging from 1 to 5; 1=0 – 10 books; 5=more than 100 books.

Home education resources - ASDHHER: three levels: 1 = low, 2 = medium, 3 = high.

Non-native parents - ASDGBRG: 1 if parents are non-native, 0 otherwise.

Father education - ASBHLEDF: Education level of father (ISCED); ranging from 0 to 7, where 1 represents no education, 2 is ISCED 2, 3 = ISCED 3, 4=ISCED 4, 5=ISCED 5B, 6=ISCED 5A and 7 = beyond ISCED 5A

Mother education - ASBHLEDM: Education level of mother (ISCED); same as father education.

Highest parental education - ASDHEDUP: same as in PISA, f=max(Father education, Mother education)

Teacher experience - ATBGTAUG: Years taught in total, ranging from 1 to 40.

Teacher female - ATBGSEX: 1 if teacher is female, 0 otherwise.

Student weight - TOTWGT: final student weight.

PD - ACBGPST1: Index of how many students are economically disadvantaged in a school. 1= 0 – 10 %, 2=11 – 25 %, 3=26 – 50 %, 4 = 50 % or more.

Total Enrollment - ACBG4ENR: Total school enrollment in fourth grade.

Instrumented class size: predicted class size from the first stage regression.

Function: $e/[floor((e-1)/25+1)]$

Average reading score: calculated by averaging over all students in one class.

Averageclasssize: averaging class size over the students in one class; since the teacher should have reported the same value for all it should be identical to the class size variable.

Class weight - WGTFAC2: class weight factor.

JKREP: contains jackknife replication information

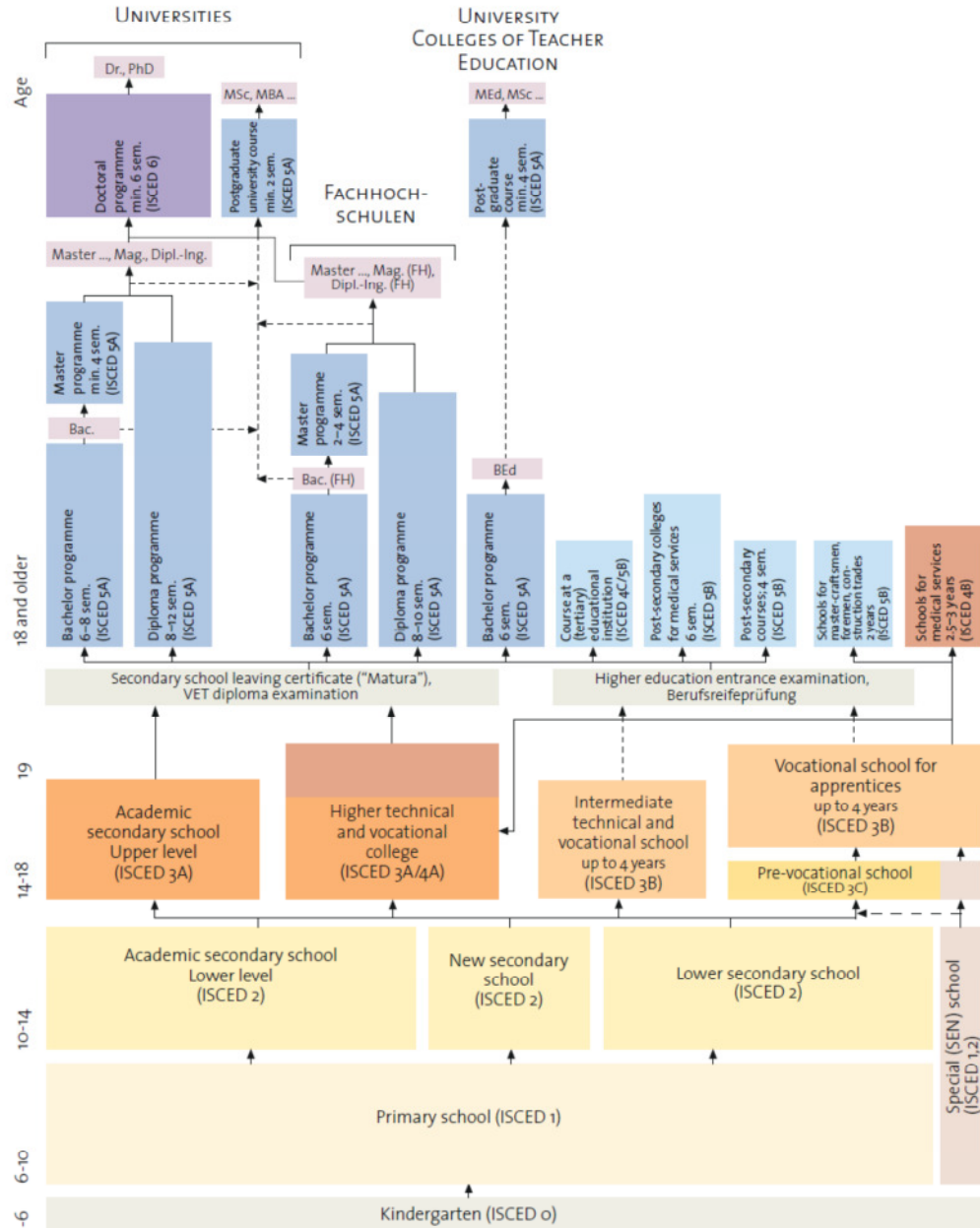**JKZONE:** contains jackknife replication information

AJKREP: averaged JKREP

AJKZONE: averaged JKZONE

AGGTOTWGT: averaged total student weights

9.3 Austrian Educational System

Austrian Educational System

ISCED: International Standard Classification of Education (UNESCO)

9.4 Stata Syntax PISA 2009

set memory 2g

use "*<replace with pisa-school-file path>*", clear

keep if CNT=="AUT"

gen totalschoolenrollment= SCHSIZE

replace totalschoolenrollment=. if totalschoolenrollment>3450

gen schoolquality=SCMATEDU

replace schoolquality=. if schoolquality>2

save"<replace with pisa-school file path>", replace

leaves dataset with a total of 282 schools

clear

Use<*replace with pisa-student-filepath* >, clear

net describe pv, from(http://fmwww.bc.edu/RePEc/bocode/p)

net install pv

installs the add on to deal with the complex PISA replicate weights structure

keep if CNT=="AUT"

gen Grade = ST01Q01

gen classsize = ST35Q01

replace classsize=. if classsize>36

gen female = 0

replace female=1 if ST04Q01==1

gen germanathome = 0

replace germanathome=. if ST19Q01>6

replaces all misreports to missing values

replace germanathome = 1 if ST19Q01==1

gen booksathome = ST22Q01

replace booksathome=. if ST22Q01>6

replaces all misreports to missing values

gen age=AGE

gen Father_ISEI=BFMJ

replace Father_ISEI=. if Father_ISEI>90

replaces all misreports to missing values

gen Mother_ISEI=BMMJ

replace Mother_ISEI=. if Mother_ISEI>90

replaces all misreports to missing values

gen wealth = WEALTH

replace wealth = . if WEALTH > 3

replaces all misreports to missing values

gen slwbp = 0

replace slwbp=. if FAMSTRUC>6

replaces all misreports to missing values

replace slwbp = 1 if FAMSTRUC==2

gen Father_education = FISCED

gen Mother_education = MISCED

replace Father_education=. if Father_education>6

replaces all misreports to missing values

replace Mother_education=. if Mother_education>6

replaces all misreports to missing values

gen highest_parental_education=max(Father_education, Mother_education)

gen home_education_resources = HEDRES

replace home_education_resources=. if HEDRES>1

replaces all misreports to missing values

gen home_possession_index=HOMEPOS

replaceh ome_possession_index=. if home_possession_index>4

replaces all misreports to missing values

sort SCHOOLID Grade classsize

by SCHOOLID Grade classsize: gen number=_n

replace number=. if number>1

by SCHOOLID Grade: egen avgcs=mean(classsize) if number==1

by SCHOOLID Grade: egen averageclasssize=max(avgcs) if classsize!=.

drop avgcs number

rename averageclasssize avgcs

generates Grade average class size which will be used as an Instrumental Variable

gen weight=W_FSTUWT

pv [weight=weight], pv(PV*READ) cmd("mean") brr rw(W_FSTR*) fays(0.5)

sum classsize female Grade germanathome booksathome age Father_ISEI Mother_ISEI slwbp highest_parental_education home_education_resources home_possession_index wealth avgcs [weight=weight]

generates the summary statistics in table 3

pv classsize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth [weight=weight], pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

generates column 1 of table 5

tab SCHOOLID, gen(sd)

generates school dummies to control for across school variation

pv classsize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth sd* [weight=weight], pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

generates column 2 of table 5

merge m:1 SCHOOLID using "<replace with pisa-school-file path>"

pv classsize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth totalschoolenrollment schoolquality [weight=weight], pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

generates column 4 of table 5

pv classsize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth sd* if totalschoolenrollment!=. & schoolquality!=. [weight=weight], pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

generates column 3 of table 5

areg classsize avgcs Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth if Grade!=avggrade [weight=weight], absorb(SCHOOLID)

generates column 1 (first stage regression) in table 9

predict yhat

rename yhat instrumentcs

pv instrumentcs Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth sd* [weight=weight], pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

generates column 2 (second stage regression) in table 9

pvclasssize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth sd* [weight=weight] if ABGROUP==1, pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

generates column 3 (robustness check regression)

rename _merge merge1

merge 1:1 StIDStd using ">path for sparte2 (identifies school type) variable "

the next five commands generate table 11; by sparte2: – command does not work with pv command

pv classsize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth sd* [weight=weight] if sparte2==1 & ABGROUP==1, pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

pv classsize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth sd* [weight=weight] if sparte2==2 & ABGROUP==1, pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

pv classsize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth sd* [weight=weight] if sparte2==3& ABGROUP==1, pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

pv classsize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth sd* [weight=weight] if sparte2==4& ABGROUP==1, pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

pv classsize Grade female germanathome booksathome age slwbp highest_parental_education home_education_resources home_possession_index wealth sd* [weight=weight] if sparte2==5& ABGROUP==1, pv(PV*READ) brr rw(W_FSTR*) fays(0.5)

9.5 Stata Syntax PIRLS 2006

use*<path of PIRLS STATA file>*

gen classsize=ATBGCSTD

replace classsize=. if classsize>30

gen female=0

replace female=1 if ITSEX==1

gen non_native_parents=0

replace non_native_parents=. if ASDGBRN==9

replace non_native_parents=1 if ASDGBRN==3

gen booksathome= ASBHBOOK

replace booksathome=. if booksathome==9

gen childbooksathome= ASBHCHBK

replace childbooksathome=. if childbooksathome==9

gen age=ASDAGE

gen father_educ=ASBHLEDF

replace father_educ=. if father_educ>7

gen mother_educ =ASBHLEDM

replace mother_educ=. if mother_educ>7

gen teacher_experience= ATBGTAUG

replace teacher_experience=. if teacher_experience>40

gen teacher_female= ATBGSEX

replace teacher_female=. if teacher_female==9

replace teacher_female=0 if teacher_female==2

gen years_teaching= ATBG4TOT

replace years_teaching=. if years_teaching==9

gen home_educ_res= ASDHHER

replace home_educ_res=. if home_educ_res==9

gen high_parental_educ=max(mother_educ, father_educ)

gen wealth=ASBHWELL

replace wealth=. if wealth==9

sum classsize female non_native_parents booksathome childbooksathome age father_educ mother_educ teacher_experience teacher_female home_educ_res high_parental_educ wealth [weight= TOTWGT]

*generates summary statistics table 4 [excluding reading score]

pv [weight=TOTWGT], pv(ASRREA0*) jkzone(JKZONE) jkrep(JKREP) jrr pirls

pv classsize female age booksathome childbooksathome home_educ_res non_native_parents high_parental_educ wealth teacher_experience teacher_female [weight=TOTWGT], pv(ASRREA0*) cmd("reg") jkzone(JKZONE) jkrep(JKREP) jrr pirls

*generates column 1 of table 12

tab IDSCHOOL, gen(sd)
gen number=(IDCLASS-IDSCHOOL*100)
sort IDSCHOOL
by IDSCHOOL: egen x =mean(number)
pv classsize female age booksathome childbooksathome home_educ_res non_native_parents high_parental_educ wealth teacher_experience teacher_female sd* if IDSCHOOL!=30 & x!=4 & x!=1 [weight=TOTWGT], pv(ASRREA0*) cmd("reg") jkzone(JKZONE) jkrep(JKREP) jrr pirls

*generates column 2 of table 12

pv classsize female age booksathome childbooksathome home_educ_res non_native_parents high_parental_educ wealth teacher_experience teacher_female if IDSCHOOL!=30 & x!=4 & x!=1 [weight=TOTWGT], pv(ASRREA0*) cmd("reg") jkzone(JKZONE) jkrep(JKREP) jrr pirls

*generates column 3 of table 12

sort IDCLASS
by IDCLASS: egen avgcs1=mean(classsize) if classsize!=.
by IDCLASS: egen AGGTOTWGT=mean(TOTWGT) if classsize!=.
by IDCLASS: egen TASRREA01=mean(ASRREA01)
by IDCLASS: egen TASRREA02=mean(ASRREA02)
by IDCLASS: egen TASRREA03=mean(ASRREA03)

by IDCLASS: egen TASRREA04=mean(ASRREA04)

by IDCLASS: egen TASRREA05=mean(ASRREA05) gen PD=ACBGPST1

replace PD=. if PD==9

by IDCLASS: gen number2=_n

gen totalenrollment=ACBG4ENR

replace totalenrollment=. if totalenrollment==9999

by IDCLASS: egen AJKZONE=mean(JKZONE)

by IDCLASS: egen AJKREP=mean(JKREP)

pv avgcs1 PD totalenrollment [weight=AGGTOTWGT] if number2==1, pv(TASRREA0*) cmd("reg") jkzone(AJKZONE) jkrep(AJKREP) jrr pirls

*generates column 1 in table 13

gen function=totalenrollment/[floor((totalenrollment-1)/25+1)]

reg avgcs1 function PD totalenrollment [weight=AGGTOTWGT] if number2==1

*generates column2 in table 13

predict instrument1

pv instrument PD totalenrollment [weight=ATOTWGT] if number2==1 & avgcs1!=., pv(TASRREA0*) cmd("reg") jkzone(AJKZONE) jkrep(AJKREP) jrr pirls

*generates column 3 in table 13

10.   References

Adams, R.J. & Wu, M.L. (editors), (2002), PISA 2000 Technical Report, OECD Publications, Paris.

Akerhielm, K. (1995), Does class size matter?, Economics of Education Review, 14(3), 229-241

Angrist, J. and V. Lavy (1999), Using Maimonides rule to estimate the effect of class size on scholastic achievement, Quarterly Journal of Economics,114(2), 533-575

Baumert, J. &Demmrich, A. (2001), Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. European Journal of Psychology of Education, 16 (3), 441-462.

Barro, R. (1997), Determinants of Economic Growth: a Cross-Country Study, MIT Press, Cambridge, U.S.A.

Betts J. R.(1998), The impact of educational standards on the level and distribution of earnings. Am Econ Rev 88(1): 266–275

Bishop, J.H. (1997), The effect of national standards and curriculum-based exams on achievement, American Economic Review, Papers and Proceedings, Vol.87, pp.260–264.

Bishop, J. H. and Wößmann, L. (2001), Institutional Effects in a Simple Model of Educational
Production, Kiel Institute for World Economics, Kiel Working Paper 1085, November (forthcoming in Education Economics).

Bologna Process Website http://www.ond.vlaanderen.be/hogeronderwijs/bologna/ (28.02.2012);

Bologna Process (newest) Budapest-Vienna Declaration on the European Higher Education Area (28.02.2012)
Link: http://www.ond.vlaanderen.be/hogeronderwijs/bologna/2010_conference/docume nts/ Budapest-Vienna_Declaration.pdf (last online: 23.04.2012)

Brunner, M., Artelt, C., Krauss, S., &Baumert , J. (2007), Coaching for the PISA test. Learning and Instruction, 17(1), 111-122.

Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des Bildungswesens (BIFIE), Schwantner U., data provider for the PISA study - school type variable

Canadian Council on Learning (2009), Lessons in Learning, Why boys don't like to read: Gender differences in reading achievement, http://www.ccl-cca.ca/pdfs/LessonsInLearning/02_18_09-E.pdf - 11.02.2012,

Card, D. and Krueger A.B. (1992), Does School Quality Matter: Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy* 100

Card, D. and Krueger A.B. (1996), Labor Market Effects of School Quality: Theory and Evidence, In Gary Burtless, editor, *The Link Between Schools, Student Achievement, and Adult Success* Washington D.C.: Brookings Institution, 1996. A longer version of this paper is contained in National Bureau of Economic Research Working Paper 5450

Card, D.(2001), Education, Earnings, and the Canadian G.I. Bill. (with Thomas Lemieux), *Canadian Journal of Economics* 34

Coleman, C. (1966), Equality of educational opportunity. (Summary report.) [Washington, D.C.] U.S. Department of Health, Education, and Welfare, Office of Education,

Denny, K. and Oppedisano V. (2010), Class size effects: evidence using a new estimation technique, Geary WP2010/51, UCD Dublin

Epstein, I. (1976), Hebrew-English Translation of the Babylonian Talmud, Baba Bathra, Volume I (London: Soncino Press).

Fertig M (2003) Educational production, endogenous peer group formation and class composition. Evidence from the PISA 2000 study. Institute for the Study of Labor (IZA), Bonn, Discussion paper no. 714

Fessler, P., Schneebaum, A. (2012). Gender and Educational Attainment across Generations in Austria." Feminist Economics, Routledge, forthcoming.

Foy, P. and Kennedy, M. A. (2008), PIRLS 2006 User Guide for the International Database Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

IEA, (PIRLS Database online), Link: http://timssandpirls.bc.edu/pirls2006/user_guide.html (last online: 23.04.2012)

Jamison E., Jamison D.T and Hanushek E.A.; (2006), The effects of education quality on income growth and mortality decline." NBER Working paper 12652.

Haider G.et al. (2001) PISA 2000. Nationaler Bericht. Österreichischer Studien-Verlag, Innsbruck.

Hanushek, E.A.(1986). The economics of schooling: production and efficiency in public schools", Journal of Economic Literature, Vol.24, pp.1141–1177.

Hanushek, E.A. (1994), Money might matter somewhere: a response to Hedges, Laine and Greenwald, Educational Researcher, 23(4), 5-8

Hanushek, E. A.(1995), Interpreting recent research on schooling in developing countries", World Bank Research Observer, Vol.10, pp.227–246.

Hanushek, E.A.(1999), The evidence on class size" in Earning and Learning: how schools matter (editors Susan E. Mayer and Paul E. Peterson) Brookings Institution Press, Washington D.C. and Russell Sage Foundation, New York

Hanushek, E.A. and Kimko, D.D.(2000). Schooling, labor-force quality, and the growth of nations, American Economic Review, Vol.90, pp.1184–1208.

Hanushek, E.A. (2003), The failure of input-based schooling policies, Economic Journal, 113 (485), 64-98

Hoxby, C. M. (1994),. Do Private Schools Provide Competition for Public Schools?", NBER Working Paper 4978, December.

Hoxby, C. M. (1996), How teachers' unions affect education production, Quarterly Journal of Economics, Vol.111, pp.671–718.

Hoxby, C.M. (1998). The Effects of Class Size and Composition on Student Achievement: New Evidence from Natural Population Variation, NBER Working Papers 6869, National Bureau of Economic Research, Inc.

Hoxby, C. M. (2000), The effect of class size on student achievement: new evidence from population variation, Quarterly Journal of Economics, 115(4), 1239-1285

Kirabo Jackson, C. (2009), Student Demographics, Teacher Sorting, and Teacher Quality: Evidence from the End of School Desegregation. Journal of Labor Economics, 27 (2): 213 DOI: 10.1086/599334

Krueger, A. (1999), Experimental Estimates of Education Production Functions" Quarterly Journal of Economics, 114(2), 497-532

Krueger, A.B. & Whitmore, D. M. (2001), The effect of attending a small class in the early grades on college test taking and middle school test results: evidence from project STAR. Econ J 111(468)

Krueger, A. (2003), Economic considerations and class size, Economic Journal, 113, 34-63

Landsberger, H.A. (1958), *Hawthorne Revisited*, Ithaca

Lazear, E.P. (2001), Educational production. Q J Econ 116(3):777–803

Leitner, A. (2001), Frauenberufe - Männerberufe: Zur Persistenz geschlechtshierarchischer Arbeitsmarktsegregation, Institut für Höhere Studien, Reihe Soziologie Nr.47, Wien.

Levin, J. (2001), For whom the reductions count? A quantile regression analysis of class size andpeer effects on scholastic achievement. Empir Econ 26:221–246

Lewbel, A. (2010), Using heteroskedasticity to identify and estimate mismeasured and endogenous regressor models, Economics Department, Boston College mimeo

Lynn, R. and Jaan, M. (2009), Sex differences in Reading Achievement, University of Ulster and University of Tartu

Macdonald K., (2008). PV: Stata module to perform estimation with plausible values, Statistical Software ComponentsS456951, Boston College Department of Economics, revised 12 Apr 2011.

Manski, C.F. (1995), Identification problems in the social sciences. Harvard University Press, Massachusetts

Martin, O. M. & Mullis, V. S. I. & Kennedy, M. A. & Foy, P. (2007), PIRLS 2006 International Report, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College

Martin O. M. & Mullis V. S. I. & Kennedy M. A. (2007), PIRLS 2006 Technical Report, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College

Ours, J.C. van (2006), Children Reading Fiction Books Because They Want To, (Center Discussion Paper, 2006-3, 2006-003) pp. 1-23

Ours, J.C. van (2008), When do children read books? *Education Economics*, *16*(4), 313-328.

Neuwirth, E. OECD (2006), PISA 2000: Sample Weight Problems in Austria. Education Working Paper No. 5

Neuwirth, E. Grossman, W. & Ponocny, I. (2006), PISA 2000 und PISA 2003: Vertiefende Analysen und Beiträge zur Methodik, Leykam, ISBN 3701175691

Organisation for Economic Co-operation and Development (2001), Knowledge and skills for life. First results from PISA 2000, Paris. http://www.pisa.oecd.org

Organisation for Economic Co-operation and Development (2002), PISA 2000 technical report, Paris. http://www.pisa.oecd.org/tech/intro.htm

Organisation for Economic Co-operation and Development (2004), Learning for tomorrow's world: first results from PISA 2003, Paris. http://www.pisa.oecd.org

Organisation for Economic Co-operation and Development (2005), PISA 2003 technical report, Paris. http://www.pisa.oecd.org

Organisation for Economic Co-operation and Development (2009), PISA 2006 technical report, Paris.http://www.pisa.oecd.org

Organisation for Economic Co-operation and Development (2009), PISA 2009 Main Study Data Management Manual, Version 2, January 2009, http://www.oecd.org/dataoecd/15/34/49023731.pdf

Organisation for Economic Co-operation and Development (2010), PISA 2009 Results: Overcoming Social Background – Equity in Learning Opportunities and Outcomes (Volume II) http://dx.doi.org/10.1787/9789264091504-en

Organisation for Economic Co-operation and Development (PISA-Information), Programme for International Student Assessment, http://www.pisa.oecd.org/pages/0,2987,en_32252351_32235731_1_1_1_1_1,00.html(last online: 23.04.2012)

Organisation for Economic Co-operation and Development, (PISA-Database 2000, 2003, 2006 2009), Questionnaires, Codebooks, Raw-Data, Compendia, Manuals, http://pisa2000.acer.edu.au/downloads.php; http://pisa2003.acer.edu.au/downloads.php; http://pisa2006.acer.edu.au/downloads.php   http://pisa2009.acer.edu.au/downloads.php; (last online: 23.04.2012)

Rice, J. (1902)**,** The public school system of the United States. New York: New York. Arno Press.

Schneeweis, **N. &Winter-Ebmer**, **R.** (2005), Peer Effects in Austrian Schools , in Empirical Economics, Vol. 32, Nummer 2-3, Seite 387-409, 2007

Schneeweis, N. & Zweimüller, M. (2011), Girls, girls, girls: Gender composition and female school choice; University of Linz, NRN Labor & Welfare State, Austria

Schulorganisationsgesetz (SchOG), considered status of legal framework: March 1[st]2012,http://www.jusline.at/Schulorganisationsgesetz_ %28SchOG %29.html; (last online: 23.04.2012)

Schwantner, U. & Schreiner, C. (2010), PISA 2009 Internationaler Vergleich von Schülerleistungen. Erste Ergebnisse. Lesen, Mathematik, Naturwissenschaft, Graz: Leykam, ISBN 978-3-7011-7744-8

Schwantner, U. & Schreiner, C. (2010), PISA 2009 Internationaler Vergleich von Schülerleistungen.Technischer Bericht.

Statistik Austria, Historical class size data, http://www.statistik.at/web_de/statistiken/bildung_und_kultur/formales_bildungswesen/ schulen_schulbesuch/index.html (last online: 23.04.2012)

Statistik Austria (2011), Bildungin Zahlen 2009/10 Schlüsselindikatoren und Analysen; ISBN 978-3-902703-77-4; http://www.statistik.at/web_de/dynamic/statistiken/bildung_und_kultur/publdetail?id=5 &listid=5&detail=461(lastonline: 23.04.2012)

Thistlewaite, D. L., Campbell D. T. (1960), Regression-Discontinuity Analysis: Analternative to theex post facto experiment, Journal of Educational Psychology 51: 309-317

Wößmann, L. and M. West (2002), Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS" European Economic Review, 50 , 695-736

Wößmann, L. (2002a), Schooling and the Quality of Human Capital, Springer, Berlin.

Wößmann, L. (2002b), Educational Production in East Asia: The Impact of Family Background and Schooling Policies on Student Performance, Mimeo, Kiel Institute for World Economics, October.

Wößmann, L.(2003), European education production functions": what makes a difference for student achievement in Europe?; N°190, European Commission, Economic Papers

Wößmann, L (2003a), Schooling resources, educational institutions and student performance: the international evidence. Oxf Bull Econ Stat 65(2):117–170

Wößmann, L. (2005), The Effect Heterogeneity of Central Exams: Evidence from TIMSS, TIMSS-Repeat and PISA, Education Economics 13 (2), 143-169

Wößmann, L. with Hanushek E. A. (2011), The Economic Benefit of Educational Reform in the European Union, CESifo Economic Studies 58 (1), 2012, 73-109 (Accepted; Working paper version available as: EENEE Analytical Report 7 (PDF))

11. Abstract

I determine the effect of socioeconomic factors, class size and differences in school types for the Austrian schooling system on reading performance. By using two large scale datasets mixed evidence on the class size effect was found. While there is no measured class size effect in the PISA study I found positive effects of smaller classes for students in primary schools. As an instrumental variable technique I used a grade average class size approach for the PISA dataset and a regression discontinuity design for the PIRLS dataset. As a consequence, if further research confirms the trend found in this paper, I suggest a further differentiation of class size between grades.

In dieser Arbeit wird der Einfluss von sozioökonomischen Aspekten, der Klassengröße und der Einfluss unterschiedlicher Schultypen auf die Leseleistung österreichischer Schüler bestimmt. Der Effekt der Klassengröße auf den Schulerfolg variiert stark zwischen den Datensätzen. Während in der PISA Studie kein Effekt festgestellt werden konnte, zeigt sich im PIRLS Datensatz ein signifikant positiver Effekt von kleineren Klassen. Als Instrumentalvariable wurden Durchschnittsklassengröße für die PISA Studie und ein Regressionsdiskontinuitätsdesign für die PIRLS Studie verwendet. Aufgrund dieser Ergebnisse wird vorgeschlagen, wenn weitere Studien den ermittelten Trend bestätigen, eine größere Differenzierung der Klassengröße zwischen Schulstufen vorzunehmen.

MICHAEL TOPF

michaeltopf@hotmail.com

Stiftsfeld 9	Phone: +43 660 2110781
4652 Fischlham	
Austria	

Educational Background

2001 - 2006	High school of tourism in Bad Ischl
2007 - 2009	Bachelor studies in economics (University of Vienna)
2009 - present	Master's program in economics (University of Vienna)
Fall term 2010	exchange semester at University of Illinois at Urbana Champaign

Work Experience

2002 - 2005	Gastronomy internships in Austria and Switzerland during summer months
2006 - 2007	Military duty in the Austrian Armed Forces for 6 months
2007 - 2008	Summer-internship at BRP-Powertrain GmbH & Co KG in Gunskirchen
2009	Summer-internship at Raiffeisen Landesbank Upper Austria (Facility Management)
Dec. 2011 – present	Trailerline GmbH

Academic Work Experience

2011 spring term	Econometrics tutorial for master students at the University of Vienna[27]
2011 fall term	Econometrics tutorial for master students at the University of Vienna[28]
July 2011 – Jan. 2012	Project Legal cartels in Austria, Christine Zulehner[29] joint with Nikolaus Fink (Austrian Federal Competition Authority), Philipp Schmidt-Dengler (University Mannheim) Konrad Stahl (University Mannheim);

Research Experience

2009	Bachelor thesis in Health Economics; Topic: The Value of Life [30]
2009	Bachelor thesis in Development Economics; Topic: Resource curse in developing countries[31]
2011 - present	Master Thesis with the working title Class size Effects and the Educational Production Function in Austrian Schools Supervisor: Prof. DI Dr. Christine Zulehner;

---

[27] Reference: Neil Foster; http://homepage.univie.ac.at/neil.foster/ ‖ e-mail: foster@wiiw.ac.at
[28] Reference: Neil Foster; http://homepage.univie.ac.at/neil.foster/ ‖ e-mail: foster@wiiw.ac.at
[29] Reference: Christine Zulehner; http://www.econ.jku.at/Christine.Zulehner/ ‖ e-mail: christine.zulehner@jku.at
[30] Supervisor: Catarina Goulão; http://sites.google.com/site/wwwgoulao/ ‖ e-mail: Catarina.Goulao@toulouse.inra.fr
[31] Supervisor: Peter Rosner; http://homepage.univie.ac.at/peter.rosner/ ‖ e-mail: peter.rosner@univie.ac.at