

Ammermüller, Andreas; Heijke, Hans; Wößmann, Ludger

Working Paper

Schooling Quality in Eastern Europe: Educational Production During Transition

Kiel Working Paper, No. 1154

Provided in Cooperation with:

Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

Suggested Citation: Ammermüller, Andreas; Heijke, Hans; Wößmann, Ludger (2003) : Schooling Quality in Eastern Europe: Educational Production During Transition, Kiel Working Paper, No. 1154, Kiel Institute for World Economics (IfW), Kiel

This Version is available at:

<https://hdl.handle.net/10419/17826>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Kiel Institute for World Economics
Duesternbrooker Weg 120
24105 Kiel (Germany)

Kiel Working Paper No. 1154

**Schooling Quality in Eastern Europe:
Educational Production During Transition**

by

Andreas Ammermüller

Hans Heijke

Ludger Wößmann

March 2003

The responsibility for the contents of the working papers rests with the author, not the Institute. Since working papers are of a preliminary nature, it may be useful to contact the author of a particular working paper about results or caveats before referring to, or quoting, a paper. Any comments on working papers should be sent directly to the author.

Schooling Quality in Eastern Europe: Educational Production During Transition*

Abstract:

We estimate educational production functions for seven Eastern European transition countries, using student-level TIMSS data for lower secondary education. The results show substantial effects of student background on educational performance and a much lower impact of resources and the institutional setting. Two different groups of countries emerge. For the first group that features high mean test scores and has progressed far in transition, large effects of family background on student performance and a higher spread of test scores illustrate the similarity to Western European schooling systems, the performance of which it surpasses. Schools of the second group produce instead a denser distribution of educational achievement, characteristic of communist societies.

Keywords: Educational production, class size, TIMSS, transition of institutions

JEL Classification: I21, P36

Andreas Ammermüller

Centre for European Economic Research (ZEW)
P.O. Box 103443
68034 Mannheim
Germany

Hans Heijke

Maastricht University
Research Centre for Education and the Labour Market (ROA)
P.O. Box 616
6200 MD Maastricht
Netherlands

Ludger Wößmann

Kiel Institute for World Economics
24100 Kiel
Germany
Phone: (+49) 431 8814-497
Fax: (+49) 431 8814-500
E-mail: woessmann@ifw.uni-kiel.de
Internet: <http://www.uni-kiel.de/ifw/staff/woessman.htm#english>

* Ammermüller gratefully acknowledges financial support from the European Commission TSER project "Education and Wage Inequality in Europe."

1. Introduction

This paper analyzes and compares the production of schooling quality in seven Eastern European transition countries striving for EU accession. A main focus of the progress in the transition countries is on reforms of institutions and a changing structure of society. The supposedly egalitarian societies in which mainly the party rank defined the social position are replaced by a new social distinction based on occupation and income. Education is the major vehicle through which the societal changes take place. The function of education shifts from keeping the social consensus of a classless society to allocating its individual members to economic roles and positions, allowing for greater differentiation by increased educational choice (Heyneman 1997). The readiness for and speed of transition depends therefore essentially on reforms in education, which prepare people for their new roles in society. Reforms include the decentralization of the educational system, which might increase its effectiveness by increasing its responsiveness to market forces at the local and national level. A greater choice of different types of institutions for students and an increased influence of parents on their children's education might as well result in higher effectiveness in the new economic terms (Heyneman 1997). The introduction of decentralization and free choice may also serve to convey the values of democracy and the market system to the population (EBRD 2000).

Furthermore, the market economies demand new abilities of students, like managing skills and high flexibility, that were not fostered under the communist regime (Berryman 2000). In the phase of transition, where old practices become obsolete and new opportunities arise quickly, allocative skills like the ability to take the appropriate decisions, which constitute an important effect of schooling (Schultz 1975), are rewarded greatly and affect the income distribution. Consistent with this reasoning, rates of return to education have been found to increase during the transition from communism to a market system (Newell and Reilly 1999; Boeri and Terrell 2002). For example, the returns to higher education more than doubled over the past ten years in the Czech Republic, as did their spread across occupations (Klazar et al. 2001). The quality and variance of educational achievement may thus have an even bigger impact on the societal structure in transition countries than in advanced countries. When the Eastern European countries gain access to the EU in the coming years and the labor markets become integrated, they need to compete with the Western European labor force. A well-educated work force is hence imperative for a successful integration.

The challenges are accompanied by new threats to the formerly high-performing education systems. Due to economic recessions in the early phase of transition, the level of finance for schools is hard to maintain. This problem relates particularly to the countries whose setbacks in the first years of transition were greater and who continue to struggle in their reform progress.

The available resources, the institutional setting of schools and especially their usage depict the quality of the schooling systems. This paper examines the impact of these factors, as well as of student characteristics and family backgrounds, on the performance of individual students by estimating educational production functions for seven Eastern European countries. While the former factors are determined by school policy, the latter display the ability of schools to diminish the impact of the environment surrounding students. By estimating production functions of many Eastern European schooling systems for the first time, this paper intends to elaborate on the significance of education in the process of transition and to give appropriate advice for improving schooling in these countries. In addition, it contributes to the widely discussed topic of the effect of resources. We also compare the characteristics and the resulting quality of the Eastern European schooling systems to those of a sample of EU member countries.

Educational production functions relate an outcome of education like educational achievement to various inputs. In this study, standardized test scores from the Third International Mathematics and Science Study (TIMSS), which are comparable across countries, are related to variables drawn from background questionnaires on student background, resources employed and the institutional setting (cf. Wößmann 2003). These unique micro-level data on test scores and background information are available on over 42,000 students in the 7th and 8th grade of the seven countries, who took the math and science tests in 1994/95. As the only measurable outcome taken into account is educational achievement, the analysis is restricted to the cognitive dimension of schooling. Other educational outcomes like civic values conveyed to students have to be neglected in the data analysis.

The main finding of the paper is the straightforward distinction that can be made between two groups of accession countries with respect to the characteristics of the schooling systems, which constitute a key part of the transition process. One group has moved decisively towards the features of Western European countries while the other cannot display successful results of transition yet. The more advanced group, consisting of the Czech and Slovak Republics, Hungary and Slovenia, outperforms most EU countries and has many traits similar to the

Western schooling systems. The schooling systems of the less advanced group, including Lithuania, Latvia and Romania, still feature characteristics of communist times and seem not yet to educate a new generation that is competitive to EU labor markets. Further findings of the paper are the relative importance of student background for explaining test scores, the ambiguous impact of resources, and the limited but existing role of the institutional setting in understanding within-country variations in test scores.

The remainder of the paper is structured as follows. The second section describes the seven transition economies and characterizes the development of their schooling systems in the 1990s. This qualitative review facilitates an assessment of the data and results and elaborates on particular characteristics that are not implied in the variables. The third section describes the TIMS study, the data for the Eastern European countries, and the model used for the estimation of the production functions. The fourth section presents and discusses the results for the Eastern European countries. The fifth section compares them to a sample of Western European countries. Finally, the sixth section concludes with a summary of the findings and an assessment of their contribution and relevance in the context of transition economies.

2. Review of the Schooling Systems in Eastern Europe

Seven countries that belong to the group of transition countries and have started talks for EU accession participated in TIMSS 1995. They include the Central and Eastern European countries Czech Republic, Hungary, Romania, Slovakia and Slovenia and the Baltic States Latvia and Lithuania. Table 1 presents information on the transition progress and level of development for each country. The figures on GDP indicate the exceptional status of Slovenia, featuring almost twice the Czech and Hungarian level of GDP per capita. Slovakia is slowly catching up to the three frontrunners. The Baltic States and Romania, the biggest country and laggard in recent years, belong to the lowest developed countries of the sample. The countries Czech Republic (CSK), Hungary (HUN), Slovakia (SLV) and Slovenia (SVN) will be referred to as the first group from now on. The second group comprises the remaining three countries Latvia (LVA), Lithuania (LTU) and Romania (ROM).

With the retreat of the communist ideology in Eastern Europe in 1990, quick and fundamental reforms of the educational systems followed. Administration of schooling was decentralized, leaving multiple authorities a say on education, including parents and the church. Additionally, national schools were established, especially in the Baltic States, to foster national culture and language. Moreover, the heterogeneity of schools increased,

changing from a system of only one basic school to more specialized institutions like the gymnasium or technical and vocational schools and in addition to private schools (cf. Filer and Münich 2000). For lower secondary education, where TIMSS took place, the choice of different types of schools is limited to one form in Latvia, Romania and Slovenia, two in the Czech Republic and Hungary, four in Slovenia and five in Lithuania. Assuming a self-selection of the more able students into more demanding types of schools with entrance examinations like the gymnasium, the spread of students with respect to ability might be greater in countries with more educational choice. However, Table 1 shows that the number of school types is not decisive for the spread of test scores that are displayed in Table A3, with Lithuania having the highest number of school types and the second lowest standard deviation. Policies also aimed at decentralization and more heterogeneity of the system within the single school types, by setting up special ability classes for both low and high performing students.

With respect to the overall variation in performance, the fast implementation of reforms in Hungary and the Czech Republic, as well as the earlier reforms in Slovenia, might have exerted a positive impact on the spread of scores. In Romania instead, where the latest educational reforms did not take place prior to 1993 and principally the communist schooling system still existed at the time the test was taken, the uniformity of the scheme might have restrained the spread of scores. As a further measure with possible influence on school quality and its variation, the average attendance rate in pre-primary education is higher in the first group of countries, where parents have to pay less and have a higher average income. The last year of pre-primary education is obligatory only in Hungary, Slovenia and Romania.

3. Data and Regression Technique

3.1 The TIMSS Data

The Third International Math and Science Study (TIMSS) was conducted in 1994/95 by the International Association for the Evaluation of Educational Achievement (IEA). Over 40 countries worldwide participated in TIMSS, making it the largest and most complex achievement study ever conducted (Gonzalez and Smith 1997). This paper considers only the sample Population II with students from the middle school years. It comprises students from two adjacent grades who have the largest proportion of 13-year-old students. They correspond to the seventh and eighth grade in lower secondary schools in the seven countries considered

here. For the analysis in Eastern European countries, data for over 42,000 students from more than 1000 schools are available.¹

The students took standardized achievement tests in both mathematics and science. The results were scaled according to an international test score with a mean of 500 and an international standard deviation of 100. In addition, each student, his teacher and principal had to complete a questionnaire giving background information on students, the community, resources of the school including teacher characteristics and the institutional setting. All available data for an individual student were merged in one database (Wößmann 2003) together with his sampling weight. Table A1 provides data on the participation at student, class and school level and the ratio of sampled students. Schools in geographically remote regions, extremely small schools and schools for students with special needs were excluded from the study, as were disabled students in regular schools. This might have led to a bias of the sampled students, especially when many disordered but not disabled students have been placed into special schools. However, all other students could be sampled and the exclusion rate was not to exceed 10 percent.

As the questionnaires handed to students and teachers comprised dozens of questions, the problem of missing data was inevitable. In order to prevent a selection bias by ignoring all observations with incomplete data and to keep the sample size high, missing data were imputed. A set of fundamental variables that are available for the greatest part of the students is selected among the explaining variables. In cases where these variables are missing, the average at the lowest level available is taken as an approximation, meaning first the class average, then school average or finally country average. Each of the other explaining variables is then regressed on this set of variables, and missing values in these other variables are substituted by the predicted value from this regression. In the case of qualitative data, the prediction was conducted on the basis of probit and ordered probit models.² For the purposes of this paper, this data imputation was conducted separately within each country.

The dataset that is built on TIMSS offers the unique opportunity to analyze and compare the educational systems in the seven transition countries. With the exception of Hungary and Poland, the earlier communist countries have not previously participated in an international cross-country study on student achievement. The dataset allows for the first estimation of educational production functions for this large group of Eastern European countries. Moreover, the quality of the available data with the immense background information on

¹ For more detailed information on TIMSS, see the internet homepage <http://timss.bc.edu>.

² For more details on the imputation method, see Wößmann (2003).

various inputs and the quantity, with data available for between 4,976 and 7,471 students in the individual countries, allow for a very thorough analysis. The sampling design to test two adjacent grades of each school also permits to estimate the between-grade variation, as will be discussed later on in the identification strategy.

An obvious limitation of each cross-sectional study is that data are only available for one point in time. It is therefore impossible to control for prior educational achievement and to consider the added value of a school year to students separately. Instead, the data compare the level of student achievement. Further possible limitations are missing variables for state or regional factors. Omitting them in the regressions might lead to a bias of coefficients. Aggregation of data above the state level is likely to exacerbate this problem (Hanushek et al. 1996), which does not apply to the micro-level database used here, though.

3.2 Description of Country Data

This section sketches the different schooling systems by considering the most noticeable mean values of the explanatory variables given in the data and used in the estimated models. The definitions and ranges of value of all 25 variables are given in Table A2, while Table A3 displays their mean values and standard deviations.

A distinction between the two groups is evident in the mean test scores that the students achieved in TIMSS. The four countries of the first group all reach mean scores in both math and science that lie above the international mean of 500. They even accomplish higher scores than most Western European nations, including Denmark, France and Germany. Czech students performed best among the participating transition countries, with average scores of 544 in math and 553 in science. The countries of the second group instead all scored below the international mean of 500, with Lithuania as the worst performer with 454 points in math, superior only to Portugal in Europe. The spread of the test scores is lowest in the low scoring Baltic States and highest for the Czech Republic, Hungary and also for Romania.

Overall, the parents of the tested students are well educated, with the slight exception of Romania. The average class size varies within the two groups of transition countries, being lowest in the Baltic States with around 22 students per class and highest in Slovakia and Romania with a class size of over 26. The share of female teachers is around 80 percent in the transition countries.

The separation between the two groups of EU accession countries holds also for the descriptive statistics of other variables. The second group has a higher share of students from

broken families, of parents belonging to the lowest educational group and a lower average education, except for Lithuania. The second group suffers more from a shortage of materials in schools and has fewer well-educated teachers. The Czech Republic and Slovakia are instead well endowed with materials and well-educated teaching staff. The schooling systems are most decentralized in Hungary and Slovenia and still very centralized in Romania.

3.3 Regression Models and Techniques

3.3.1 The Basic Specification

To estimate educational production functions for the individual countries at the student level, the following general model will be employed:

$$(1) \quad T_{ics} = \beta_0 + B_{ics}\beta_1 + R_{cs}\beta_2 + I_{cs}\beta_3 + v_s + \varepsilon_{ics}$$

where T is the math or science test score of student i in class c at school s , B is a set of background variables on the student and her family, R are measures of resources used and teacher characteristics, I is a set of variables reflecting the institutional setting and v and ε are error terms at the school and student level, respectively.³

The survey design of TIMSS demands specific regression techniques for the estimation of the educational production functions. The sampling design of TIMSS contains both varying sampling probabilities for students from different schools and clustered data. Giving different weights to students who had different sampling probabilities allows obtaining nationally representative coefficient estimates. This is done by applying weighted least squares (WLS) as a regression technique for all regressions performed with the data (cf. Wooldridge 2001).

The second issue of clustered data is more troublesome. In each country, participating schools were chosen in a first step, and then the classes which took the standardized tests within each school. Therefore, the primary sampling units (PSU) are not the individual observations, the students, but instead their schools. The problem arises that the observations within the cluster of a school are not independent as they share common characteristics, which cannot be totally controlled for. The error term of the regression may therefore be more complex than assumed by conventional least-squares methods, comprising besides an

³ The information on individual student test scores and background variables provided by the TIMSS micro-level database allows for a more precise estimation of coefficients and less bias than does aggregated data used for most estimations (Card and Krueger 1996). Although also using the TIMSS database, Hanushek and Luque (2002) do not employ it at the student level but aggregate it to the classroom level instead.

individual component also class and school elements. Ignoring these latter parts can lead to spurious regression results, as the supposedly independent observations depend on each other.

The method of clustering-robust linear regression (CRLR) offers a solution to the obstacle. It allows for any dependence of observations within the PSU, demanding only independence across PSUs. Ignoring the cluster design and treating dependent observations as independent by using standard formulas would result in standard errors that are too small (White 1980). There is no impact on the coefficients of the parameters, however. Therefore, the standard errors estimated by OLS, or in our case WLS, need to be corrected. The clustering-robust variance-covariance matrix of the coefficient estimates $\widehat{V}(\widehat{\beta})$ can be approached by

$$(2) \quad \widehat{V}(\widehat{\beta}) = (X'WX)^{-1} \left(\sum_s X_s' W_s e_s e_s' W_s X_s \right) (X'WX)^{-1},$$

where $\widehat{\beta}$ represents the coefficients β_0 to β_3 from model (1), the matrix X all explanatory variables, W is the weight matrix and $e_s e_s'$ are the cluster matrices of the WLS residuals from each cluster e_s (White 1980: 821; see also Deaton 1997: 73-78).

This formula offers a consistent estimate of the variance-covariance matrix of the WLS estimator, even if the error variances differ across clusters and arbitrary correlation patterns exist within clusters. A supposition is a fixed cluster size as the sample size increases, which is fulfilled because the number of students tested in each school is independent of the overall number of students in the sample. For the estimation of the educational production functions of the individual countries, model (1) will be estimated by CRLR. Hence, the WLS coefficients and clustering-robust estimated standard errors will be presented for the production functions. Whenever merged data of several countries are analyzed, a dummy for each but one country is included in the regressions. This allows for a correlation of error terms within countries, which is likely.

3.3.2 Estimating the Effect of Resources

The impact of a school's resource endowment on its students' educational achievement is a hotly disputed topic in the literature on educational production (cf. Hanushek 2003; Krueger 2003). The causality of the resource-performance link is ambiguous since the supposedly exogenous resource variables might be influenced by actual performance of the students and might thus be endogenous. Estimating the effect of resources, especially of class size, on student achievement is therefore burdensome. Inasmuch as the TIMSS data come from an observational study and not an experiment, the coefficient may be biased by the non-random allocation of students to different class sizes, both between schools and within schools.

Controlling for these biases is essential in order to obtain a consistent estimate for the effect of resources.

Several mechanisms are imaginable that lead to the non-random allocation of students between schools, depending on the schooling system. Parents may either move to districts that offer smaller classes to their children, or the local school administration might put worse performing students into schools with smaller classes, especially when various types of schools are available. In both cases, between-school sorting takes place and biases the estimator of the class-size effect.

One strategy to eliminate all variation between schools is to control for school fixed effects. For its implementation, a dummy variable D for each school is included in model (1), leading to the function

$$(3) \quad T_{ics} = \alpha D_s + \beta_0 + B_{ics} \beta_1 + R_{cs} \beta_2 + \varepsilon_{ics}.$$

The institutional variables I that are mostly determined at the school level, are not included in this model because the inclusion of the school dummies removes all possible variation between them. This model is referred to as school fixed effects (SFE) model.

Having controlled for between-school variation and having only between grade variation left, a potential bias may still stem from within-school sorting. In order to account for the non-random allocation of students to different class sizes within a school, the technique of instrumental variables (IV) is used. Akerhielm (1995) instrumented actual class size with average class size for a given subject in the school and student enrollment at the given grade. The legitimacy of using student enrollment as an instrument is questionable, however, as overall school enrollment may also exert an impact on student achievement (Summers and Wolfe 1975; Angrist and Lavy 1999).

In this analysis, actual class size is instrumented by average class size at the grade level. It is highly correlated with actual class but not with the error term. It affects student achievement only indirectly through the impact on actual class size. For the regression, a two-stage least squares estimation is used. Actual class size S is regressed on average class size at the grade level A , the other exogenous variables Ψ and the school dummies D :

$$(4) \quad S_c = \alpha_1 A_c + \alpha_2 \Psi_{ics} + \alpha_3 D_s + \varepsilon_{ics}.$$

The predicted value $\widehat{S}_c = S_c - \varepsilon_{ics}$ consists of the non-random part S_c and the random part ε_{ics} . Using only the systematic part of S_c , no correlation will exist between \widehat{S}_c and ε_{ics} , allowing the second stage of the regression to produce a consistent estimator for class size:

$$(5) \quad T_{ics} = \gamma_1 \widehat{S}_c + \gamma_2 \Psi_{ics} + \gamma_3 D_s + \varepsilon_{ics}.$$

Having controlled for between- and within-school sorting effects, the coefficient γ_1 is then an estimate of the causal relationship between class size and student achievement. Model (5) combines the school fixed effects model and the instrumental variable technique and is referred to as SFE+IV.⁴ The only variation left to explain is within-school between-grade variation. Therefore, comparable data are needed for at least two grades, if possible adjacent ones, for each school to implement this estimation strategy. Fortunately, the TIMSS data fulfill this requirement and thus the regression strategy can be implemented.

4. Results

The results of the estimation of educational production functions for the seven transition countries are discussed separately for each category of explanatory variables. In addition, a closer look is paid to possible interaction effects and to the effect of resources that are measured in the form of class size.

4.1 Results of the Educational Production Functions

The production functions have all been estimated by CRLR (cf. section 3.3.1) using model (1). First, only the student background variables are regressed on the test scores, then the resource and the institutional variables are added to the production function, respectively.

4.1.1 Student Background

The variables of the student background in Table 2 feature the largest and most significant coefficients of the production function, which is in line with the results from other estimations in the literature. The coefficient for ‘Upper grade’ is the largest one of the function in absolute terms and varies between 87 in LTU and only 34 in ROM. It is statistically significant at the one percent level in all countries and has a positive sign as could be expected. The coefficient indicates the value added at the eighth grade. In general, it is higher in the countries that have

⁴ See Wößmann and West (2002) for details on the SFE+IV method.

a higher mean test score, with the exception of LTU. The effect of age is negative because we control for the grade level. A one-year-older student, all else equal, scores on average more than 33 points less in CSK, which is a third of the international standard deviation of the test scores. In ROM instead, the difference is only 7 points. Female students score on average lower than their male counterparts in the subjects of math and science. Only in LTU for math, they score significantly better at the 10 percent level.

Students who were not born in the respective country could be expected to perform worse. For math, this is only statistically significantly the case in CSK, whereas the coefficient is positive and statistically significant at the 5 percent level in ROM and statistically insignificant in the other countries. For science, statistically significantly negative coefficients are reported for CSK, HUN, SVN, LTU and LVA. The general trend in the countries implies that immigrant students are still disadvantaged, with the exception of ROM. There, the enormous effort of integration and their special rights like the provision of text books in minority languages seem to have fostered their educational achievement more than that of other students. In LVA, which has a huge share of immigrated Russian students, the integration is not successful and might partly explain the low average score of LVA students. In CSK and HUN, students living with both parents performed statistically significantly better than students not living with both parents.

The highest level of education achieved by the students' parents is of great importance. The effect of completed secondary education is highest in CSK with 24 points difference to students of parents with some or no secondary education in the case of math. However, it is statistically insignificant in HUN, LTU, LVA and ROM for the science estimates. This may be due to the high share of originally missing values, which have been imputed. In the four countries, over a fifth of the observations for the parents' educational level are missing. In LTU, over 40 percent are imputed. The impact of finished university education is even higher, up to 49 points in SVN for math, and is statistically significant at the one percent level in all countries except for ROM for science. In ROM, the difference to the reference group is lowest with only 21 points, and it is rather low in LTU and LVA as well.

The variable 'Books at home', which is included in its logarithmic form, has a statistically significantly positive coefficient at the one percent level in each country.⁵ Its value is greatest in CSK and HUN, where a one-unit increase on the natural logarithmic scale, equivalent to a

⁵ Using the natural logarithmic form for 'Books at home,' as well as of 'Teacher experience,' makes the production function non-homothetic. This implies that the marginal effect of particular variables depends on their level or the scale of production. Figlio (1999) illustrates that this functional form is more likely to grasp the 'true' effect of resources.

multiplication of the value by 2.7, improves the test score by 23 points. For ROM, the value is almost half of it, only 12 points. In LVA and LTU, the values are also rather low, whereas the other countries are closer to the Czech and Hungarian level. The variable has a huge influence on the test score. All else equal, a student whose parents have around 200 books at home scores over 100 points higher in math compared to a student with hardly any books at home, in the countries of the first group.

The location of the community has a statistically significant impact only in HUN and ROM, where students who live close to the center of a town score up to 20 points higher than students from non-urban areas. The community location should make a positive difference in countries with a low accessibility of schools in rural areas. In ROM, this has been described as a problem of the schooling system. Moreover, other effects of group pressure that depend on the community type and are not included in other variables may play a role.

The coefficients from the regressions for the science test score are very close to the math estimates but on average slightly lower. This suggests that student characteristics and background have less effect on the science test score, so that it depends on a student's 'quality' to a lesser degree. Probably the students' interest in the subject, which is not captured by the included variables, makes a greater difference in science. The only coefficient that is consistently larger in absolute terms for science is the one on students' sex. The negative effect of being a female student is even greater in science than in math. In CSK, this effect almost offsets half the effect of being in a higher grade, which makes a huge difference. Female students' enthusiasm is thus even lower for science than for math, if we do not expect them to be less capable, which would be a cultural effect.

Across the different countries, a pattern of the magnitude of the effects is apparent. In CSK and HUN, the coefficients always have the greatest values in absolute terms. In ROM and LVA instead, the values are the lowest whenever they are statistically significant, except for the community location. In general, the countries belonging to the first group have higher coefficients than the countries of the second group. This pattern is especially clear for the variables concerning the students' family background. Table 3 illustrates the impact that the difference between the worst and best scenario of possible student characteristics and backgrounds has on the math and science test scores. For the worst scenario, students are deemed to be one year older, female, immigrants, not living with both parents, their parents have no or only some secondary education, about five books at home and live in a rural area. This scenario is compared to the best possible one: younger, male, native, living together with both parents, parents having finished university education, about 250 books at home and

living close to the center of a town. Only the coefficients that are statistically significant at the five percent level are considered for the computation reported in Table 3.

The numbers underline both the importance of the background variables for test score differences and the reported distinction between the first and the second group of countries. The impact of 219 points in CSK for math makes up for almost half the test score mean of the country. Furthermore, the effect is lower for science than for math, except for LVA. The lower coefficients in the second group of countries imply that background differences between students affect the test scores less in this group. This suggests that in these countries, where the reform process commenced later and where, at least in the Baltic States, the Russian grip over the country was strongest, education's role of producing an egalitarian society is still stronger marked than in the first group of countries. The outcomes are relatively low scores, lower variations in the test scores in these three countries and lower returns to individual characteristics in schools. The extreme two cases are ROM for a system that seems hardly unchanged from communist times and CSK with great returns to individual background features.

4.1.2 Resources and Teacher Characteristics

The coefficients of the category of resource and teacher inputs are shown in Table 4 and indicate the effects that differences in school endowments and in the teaching staff have on their students' test scores, given the assumption that these endowments are exogenous to student performance. The estimated effects are useful for policy implications, as resources are allocated to schools by policy makers. Class size was chosen as the primary resource variable. Whenever it is highly statistically significant, it is positive. This is the case in CSK, LTU and ROM for math and in CSK for science. Only in LVA, the coefficient of science class size is negative and statistically significant at the 10 percent level. However, half of the respective values are originally missing in LVA. The positive signs could be expected from prior studies but are counter-intuitive. An increase in the math class by one student in CSK is related to a 2.23 points improvement of the students' scores, all else equal. The policy implication of these least-squares estimates would be to raise class sizes and employ fewer teachers, which seems to conflict with common sense. However, as this result might be driven by a simultaneity bias, the evidence on class size is examined in greater detail in section 4.3.

The specific needs of schools are reflected in the variable 'Great shortage of materials', as compared to some or no shortage. It is mostly negative, though it is statistically significant only in CSK and LTU, whereas it is positive in LVA.

The characteristics of teachers give only small insights into a further explanation of students' test scores. Teachers' sex has no statistically significant influence at all and the direction of a possible effect is not consistent across countries or between math and science. The high share of female teachers in the transition countries thus does not seem to have an impact on their average scores. The experience of teachers instead has a statistically significant impact in some countries. For math, positive effects are estimated in HUN and SLV and for science in ROM. The latter one has the highest value of 10 points per one-unit increase on the natural logarithmic scale, which corresponds roughly to a tripling of experience. Students with a science teacher who has 7.4 (20) years of experience score 10 points better than students with a teacher with only 2.7 (7.4) years of experience. In science in CSK, the effect of experience is negative, although not statistically significant. This might be partly due to the high average age of teachers in the country, which is the highest of the sample. From a certain age on, additional experience might be outweighed by the age of the teacher. If this is the case for a large portion of teachers, the effect might turn the sign of experience even negative. Then the effect of experience would rather be reflected by a negative quadratic function than a logarithmic one. However, testing the function showed that it suited the data not as well as the logarithmic form in most countries.

The educational level of teachers has ambiguous effects on students, which might depend on the low variation between the educational levels. These are mostly determined by the state. For science, the coefficient on teachers' having a Bachelor degree is statistically significantly positive in SVN and statistically significantly negative in LTU. The highest educational level, which is equivalent to a Master or Ph.D., has a statistically significant positive effect in ROM for math and in CSK and ROM for science. For HUN and SLV, the sign is negative but not very informative due to the low variation. The large positive coefficient for the master level for math in ROM implies a huge difference of 52 points between students' scores only due to their teacher's education. However, as less than one percent of teachers in ROM preside of a master degree, the estimate may not be representative.

The coefficients for resources and teacher characteristics are far less statistically significant and of a lower magnitude than those for the first category of student background variables. Moreover, the effects of greater endowments of schools are somewhat ambiguous. Most unexpected signs of effects can be attributed to the unusual distribution within countries, or are discussed more extensively later on in the case of class size. However, there is no consistent picture that clearly indicates the merits for students from greater resources or better staff. A proper endowment with materials, teachers' experience and their educational level of

a Master's degree still seem to be related to higher test scores of students in some cases. Possible beneficial side effects of the variables that are not grasped directly by the variable itself are examined later on in section 4.4 with the help of interaction variables.

4.1.3 Institutional Setting

Most of the variables that describe the institutional setting at the school level through stating the degree of responsibility and autonomy that school heads and teachers have are statistically insignificantly related to student performance. Table 5 presents the results. Concerning school responsibility, autonomy of schools over the budget, supplies and hiring teachers has a statistically significant coefficient only in HUN for math, where it is positive and of a high magnitude of almost 11 points for each additional field of autonomy. The positive effect may be attributable to the incentives imposed upon schools and teachers by central exams and curriculum, which are in place in all countries except for LVA and for science in ROM. When schools have the freedom to determine teacher salary instead, this has a statistically significantly negative coefficient in LTU for math and in HUN for science. The incentive for opportunistic behavior seems to be too high because it concerns foremost the benefit of teachers. Although the coefficients are imprecisely estimated, CSK may be judged to feature the best incentive system, as all its coefficients of decentralization are positive.

When teachers have a strong influence on the curriculum, the direction of the effect seems to depend on whether they act individually or collectively. In the former case, the coefficient is mostly positive, and the only statistically significant coefficient in CSK for math implies a difference of 14 points to the setting when individual teachers have no influence on the curriculum. The effect of a collective influence is instead negative in the majority of cases but never statistically significant. Individual or class teachers have an informational advantage and do not act as an interest group, which is the case for collective teachers' influence or that of teacher unions.

The possible impact of the subject teachers' strong influence on specific matters is estimated by two variables. The first is about the money for supplies and subject matter, which is negative in the majority of countries. However, there is only one statistically significant coefficient in SVN for math, which is highly positive. The coefficient on the second variable on 'Kind of supplies or textbooks' is mostly positive, but statistically significant only in ROM for math, and there is a statistically significantly negative coefficient in HUN for science, although at a lower significance level.

The two further variables on the institutional setting, ‘Homework’ and ‘Uninterested parents limit teaching’, always have the expected signs (positive and negative, respectively) when they are statistically significant. For the former, this is the case in LVA and ROM for science and in CSK for math. For the latter, it is the case in CSK and LVA for math and in CSK and SVN for science. When the teacher notices that parents’ unwillingness to cooperate limits his teaching, students in CSK score on average around 20 points less than if not. An additional hour of average homework per week for math increases test scores by 19 points in CSK, all else equal. In HUN, LTU, SLV and SVN, homework has no statistically significant influence on either math or science scores.

A major shortcoming of the results for the effects of the institutional setting on students’ test scores is the weak statistical significance of the coefficients. They point in the expected direction in the majority of cases and are often of a considerable magnitude, but no unambiguous conclusion can be drawn due to their high standard errors. The limited number of schools, around 150 per country, allows only for little measurable variation within a country and leads to low degrees of freedom when all 23 explaining variables are included in the regression. This concerns the variables on school responsibility and teacher influence on the curriculum, which are measured at the school level. For the other variables, which are measured at the class level, the degrees of freedom should suffice but the effects are not clear-cut either. It seems that differences in the institutional setting are mainly relevant for understanding the cross-country variation in student performance (Wößmann 2003), whereas the descriptive statistics in Table A3 show that there is hardly any variation in the institutional setting within most Eastern European countries which could lead to a variation in test scores. Still, if we compare the optimal institutional setting with the least favorable one, taking into account only coefficients that are statistically significant at least at the ten percent level, students score on average around 50 points higher in CSK for math and around 20 points for science. A difference of 50 points is half of the international standard deviation of test scores and around a tenth of the mean score. Hence, institutions do matter in some countries, and their setting should not be neglected, especially since a modification might be achievable at lower cost than in the other categories.

4.1.4 Explanatory Power of the Three Categories of Variables

The statistical significance and magnitude of the coefficients from the three categories of variables are reflected in their contribution to the explanatory power of the models as seen in the values of R^2 in Table 6. The 9 student background variables explain the greatest share of

the variation in test scores. The contributions of the other two categories, resources and institutional settings, which comprise 6 and 8 variables, respectively, are only supplementary and comparable in size to each other. Table 6 reflects only one order in which the categories appear in the function, which can influence the R^2 . Altering the order of appearance changes little, however. The highest R^2 for the entire model is reached in LTU, where a quarter of the variation of the science test scores can be explained by the production function. ROM instead features the lowest values in both math and science, which is only slightly above 11 percent. The values differ greatly between countries, which might either suggest that the model used to estimate the production function suits certain countries better or that the quality of the data is lower in others.

4.1.5 Sensitivity Analysis of the Imputation Technique

The estimations of the production functions have also been performed with the original values only. For each variable included in model (1), all imputed values (cf. section 3.1) have been dropped. These ‘robust’ estimates are not affected by the method of imputation. This sensitivity analysis of the results examines whether changes in the data, in this case the introduction of changes through the imputation of values, alter the outcome of the regressions, in which case inferences from these data would seem fragile (Mukherjee et al. 1998).

In general, there are no great distinctions between the two differently estimated coefficients for each variable. No statistically significant variable changes its sign. However, some statistically insignificant variables change signs and for others the statistical significance level changes. This is often the case for variables that belong to the set of fundamental variables and have over 20 percent of missing values. Then, just the average of the next highest level is taken as an approximation. In HUN and LTU this can be observed for the coefficients of the educational level of parents, which are always less statistically significant with the imputed values included. For variables that do not belong to the fundamental set, the imputation technique seems to be more efficient. They are imputed by taking the expected values from a previous regression on the set of fundamental variables (cf. section 3.1). The coefficients for the effect of science class in LVA, where more than half of the values are missing, are almost identical.

Given that the estimates are so close to each other, the imputation technique should not have led to a bias of the data and the inferences are not fragile. Having the full dataset available for the estimations is a great advantage because the higher amount of observations allows for a better explanation of the variation. This is shown by the decreasing standard

errors. Basing the estimation of the effect of school or class differences on even fewer observations might have led to even weaker inferences. Thus, the imputation of missing values was a worthwhile step in the estimation of the educational production functions.

4.2 Interaction Effects

This section considers possible interaction effects between variables of the educational production function. The interaction terms have been included separately at the end of model (1) and indicate any further effects a variable may have in connection with other variables.⁶

In both math and science, students seem to perform better under a teacher of the same sex. This is true for both boys and girls. The additional effect of teacher experience of teachers who hold a Master degree on students varies across countries. It is twice statistically significantly negative in ROM, due to the extremely high coefficient of the Master level. In HUN, the effect is statistically significantly positive for science, where the coefficient of a Master degree is statistically significantly negative, and negative for math, where the Master coefficient is positive. The interaction effects seem thus to offset some of the counterintuitive effects in the production function. The effect of class size does not consistently depend on the experience of teachers or on their educational level. When teachers hold a Master degree, class size exerts an additional significantly negative effect on student performance in three cases and a significantly positive effect in two cases.

When the students' parents have more than 200 books, the positive coefficient of class size is intensified in five countries. This might be a further proof of the selection of low performing students to smaller classes and consequently of better performing students to rather larger classes, if we assume that the number of books at home is a proxy for student ability but does not fully capture the effect. The positive interaction term is in line with the hypothesis of Lazear (2001) that the optimal class size for good students is larger than that for bad students, who disrupt learning in the classroom.

4.3 Analyses of the Class-Size Effect

In order to give more scrutiny to the possible endogeneity bias in the least-squares estimation of resource effects, we analyze the class-size effect in greater detail. We first look at class-size effects in different segments of class sizes to see whether this is indicative of sorting of students into differently sized classes, and then we implement the model combining school

⁶ Detailed results are available from the authors upon request.

fixed effects and instrumental variables derived in section 3.3.2 to eliminate any effects of between- and within-school sorting from the estimate of the class-size effect.

4.3.1 Class-Size Effects for Class Segments

In each country, the class-size effect is estimated separately for three segments of class sizes, the lower, middle and upper segment. Each segment in a country is constructed to contain approximately the same number of students, so that the segments cover different class sizes across the countries. In CSK (HUN) for example, the lower segment comprises classes up to a class size of 23 (21), the middle one classes between 24 (22) and 26 (25) and the upper classes with more than 26 (25) students. In order to avoid further bias, all explanatory variables of the production function are included in the estimation and control for other effects. As the number of classes whose size is estimated is greatly reduced in the regressions, standard errors of the class-size coefficient increase for the individual categories.

It is enticing to compare the overall coefficients of class size to the coefficients from the segments. If the coefficients for class size of the segmented student population were mainly consistent and pointed in the same direction as the overall coefficient, this would give evidence for no selection bias. Then the effect would be identical for different ranges of class sizes. However, if there were greatly varying coefficients and especially statistically significantly negative ones among them, this would point to an outside involvement like the non-random allocation of students. If this were the case and the low-performing students were put into smaller classes, then the effect for a limited range within which no selection takes place might still be negative. If the segment coefficients were negative but the overall coefficient positive, a selection between segments would seem likely. A positive coefficient would indicate that selection takes place within the considered range instead.

The results shown in Table 7 cannot reveal any clear evidence on whether between- or within-school selection takes place, though. The results from the segmented class-size estimations are not very consistent. For each country, both positive and negative coefficients are reported, with the exception of LVA with only negative coefficients. There are 11 statistically significantly positive coefficients versus 4 statistically significantly negative coefficients. For the merged dataset EAST that contains all seven transition countries and country dummies, there are statistically significantly positive coefficients for the overall estimates and statistically significantly positive ones for math in the middle and upper segment. The variation of the coefficients is greatest in the middle segment and lowest in the lower segment. For science more negative coefficients are estimated as compared to math,

where only the LVA coefficient in the upper segment is statistically significantly negative. The inconsistency of the coefficients across countries and across segments can be interpreted as an indication of a possible bias of the class-size coefficient in the production function but is clearly no proof of it.

4.3.2 Eliminating Effects of Between- and Within-School Selection

In order to control for selection that takes place between schools when measuring the class-size effect, a dummy for each school in the country is added to the regression, leading to the SFE model (3) derived in section 3.3.2. The few schools that tested only one class are excluded from the estimation of this model. The class-size coefficients of the SFE model shown in Table 8 are smaller than for the previous regression in all countries except for LTU and ROM. The standard errors increase only slightly in some cases. The results indicate that when excluding the effect of between-school selection of students, a positive effect of class size on test scores appears less likely. The change is most drastic in CSK, where the statistically significantly positive class-size coefficient turns negative for math and is around zero for science. In LVA for math, a statistically significantly negative coefficient results from the control for school fixed effects. The only positive remaining coefficients are in LTU and ROM for math and in ROM for science. In these two countries, the different model had hardly any influence on the coefficients.

In a second step, the selection within schools is additionally controlled for by using the average class size at the grade level as an instrument for actual class size (cf. section 3.3.2). The model and corresponding technique (SFE+IV) can only be applied to three countries, for which sufficient data are available on the instrument. For math, the consistent model leads to slightly positive coefficients in CSK and SVN, which are statistically highly insignificant, though. In ROM, the former significantly positive coefficient turns negative. This is as well the case for science, where all three coefficients for class size are negative, however still insignificantly. The estimates correspond to the intuitive reasoning that a smaller classroom is a better learning environment for students and should benefit their scores. The relatively large standard errors of the coefficients do not allow for pinpointing the exact effect. It can only be said that it is approximately close to zero or even slightly negative. The results do not support a positive resource-performance link, but they do show that the counterintuitive least-squares coefficients are likely to be biased.

In CSK, selection between schools seems to be the major cause of bias in the class-size coefficient. Low performing students seem to be allocated to schools that have lower class

sizes than those schools of the better students. In ROM instead, between-school selection seems to introduce no bias, but within-school selection does seem to. Low performing students are not selectively allocated to certain schools but to certain classes of lower size within schools. This might as well be the case in LTU and could explain the statistically significantly positive coefficient of class size after controlling for SFE. The difference in the origin of the selection bias seems to be related to the structure of the school system. In the more diverse systems in CSK and HUN, students of the considered age of about 13 can choose between two types of schools, and the more able students are allocated into the gymnasium. In ROM instead, only the general school exists and students cannot be selected. Therefore, if low performing students are to receive more resources, any selection has to take place within schools.

5 Comparison to Western European Countries

In order to see how far the different Eastern European school systems have converged towards the Western European ones, we compare the educational production functions of Eastern and Western European countries. The merged functions of all Eastern European countries (EAST), of the first group of countries (FIRST) and of the second group of countries (SECOND) are compared to a sample of Western European countries (WEST) that includes Austria, Denmark, France and Germany. They all are central European countries and long time members of the EU and should hence be well suited for a comparison. The respective country dummies are added to the regressions.

If the aforementioned assumption is true that the schooling systems of the first group of countries have already moved significantly towards those of democratic market economies, then the production function of the Western European sample should resemble more the function of the first group than that of the second group. Table 9 presents the coefficients of estimating model (1) with the math scores as the dependent variable. The mean math scores in the bottom right indicate that the average performance of the sample of Western European countries is only slightly above the average for Eastern Europe and right between the first and the second group of countries. The coefficients for WEST are clearly closer to the first group and are often opposed to the estimates of the second group. In the following, we compare the estimates of the Western European countries mainly to those of the first group and refer to the mean values of the explanatory variables.

In the category of student characteristics and background, the coefficients of WEST and FIRST are very close to each other, except for ‘immigrant’, where the effect in WEST is clearly negative. The integration of immigrant students in schools seems to succeed better in the former communist countries, which were accustomed to integrate workers and their families from other communist countries. Moreover, the Western countries have higher shares of immigrated students, which may complicate integration. The average education of students’ parents is lower in WEST than in both Eastern groups. The number of books at home in WEST is instead slightly higher than in the second group but well below that of the first group. It is obvious that returns to the individual characteristics and the effects of student background are higher in WEST and the first Eastern group than in the second Eastern group. In the latter, education has not fully taken over the role of differentiation of students for their later work in the labor market. Instead, schools still seem to be supposed to diminish the impact of student characteristics and produce a homogenous output of students. The lower standard deviation of the math test score of the second group reflects this focus of educational policy. However, the countries of the first group might have overshot their mark because their standard deviations are already above that of the Western European countries. Allowing for too much deviation between students’ performance might result in a great gap between social groups and pose a threat to the social consensus.

The positive class-size coefficient is even bigger in Western Europe than in the two groups that have a higher average class size. A shortage of materials is reported rarely in WEST and the first group but by every third school of the second group. Its positive effect on test scores in WEST is counterintuitive. The impact of teacher characteristics and education is very similar, except for the estimate of the highest educational level of teachers. In WEST, no estimate apart from class size and teachers’ sex is statistically significantly different from zero. Moreover, both coefficients of teachers’ education are opposed to the expected results. The difference between the 40 percent share of female teachers in WEST and that of over 80 percent in EAST is striking. The share of teachers that obtained a Bachelor degree is highest in the second group and for a Master degree in the first group, while the Western countries have mean values in between the two groups.

The estimates of the category on the institutional setting of the Western sample are again consistent with the expected effects and closer to those of the first group. On average, schools in WEST have a lower degree of autonomy than those of the first group, where it is highest, and even those of the second group. However, the influence of teachers is higher in Western countries. The distinction between the positive effect of individual teachers’ and the negative

effect of collective teachers' influence on the curriculum is especially apparent. The highly statistically significant impact of homework and uninterested parents is also striking.

The findings of the comparison are consistent with the prior assumption that the schooling systems of the Western European sample are closer to the first group of Eastern European countries. In several aspects like the returns to individual student characteristics and the deviation of test scores, the first group seems to have surpassed the Western countries already. It might be hypothesized that the differentiation of students may decrease again in the first group in the following years, when the biggest transitional steps have been achieved and an equilibrium between the new market forces and the need for social cohesion has been found. The equalization of educational outcomes is characteristic for the second group instead and is illustrated by the low effects of student characteristics on their performance and the lower deviation of test scores.

6. Conclusion

The analysis of the schooling systems of seven Eastern European countries by means of estimating their educational production functions reveals several distinctive features. First, the countries can be divided into two groups, which share similar characteristics in their economic development, the properties of their schooling systems and the effects that the various factors have on student test scores. The first group of countries, which includes the Czech Republic, Hungary, Slovakia and Slovenia, commenced the political and economic reforms earlier than the second group and features a higher level of both political and economic development. In the second group, the Baltic States instead remained under tight Russian control until 1991 and suffered great economic damage from this strong link. Romania's political struggles delayed any reforms and turned it into the political and economic laggard of this sample of Eastern European countries. Concerning their schooling systems, the average student of the first group is endowed with more favorable characteristics and a higher level of directly measurable resources at schools, especially in the two countries of the former Czechoslovakia. Further, the first group features the two most decentralized schooling systems with Hungary and Slovenia.

When regarding the impact that the individual factors have on student performance, as estimated by the educational production function, distinct patterns emerge for both groups. The effects of the very significant variables on student background, especially of family circumstances, are consistently higher in the first group, which introduced reforms to the

education systems earlier and seems to have changed the role of education away from forming an egalitarian society into a mode of differentiation for the labor market. The comparison to the sample of Western European countries underlines the move that the first group has made away from communist to Western schooling systems during the first four to five years of transition. It even surpasses the Western sample in several specific features, such as the high returns to individual characteristics of students and a higher variation in test scores. In the second group, returns to individual characteristics and especially to family background are instead lower and students' performance varies less. This gives students more equal opportunities irrespective of their background but seems to be associated with a lower average performance. Thus, the second group still features the patterns of a schooling system whose primary role is to produce a homogenous output of students.

The coefficients for resources and the institutional setting are less significant and of a much lower magnitude in all countries and not consistent within groups. The merit of increased resources is illustrated for the experience and educational level of teachers in most countries. The effect of class size remains ambiguous, but positive coefficients were shown to be biased by student selection. Thus, it seems more likely that there is no effect or even a slightly negative one, especially for lower levels of class sizes. As the reduction of class size is very costly, it is doubtful whether a minor negative effect of larger classes justifies the enormous effort of decreasing class sizes. Although the institutional setting has comparatively little impact compared to student background, differences in the autonomy of schools and teachers are shown to matter in some countries and are relatively effortless to modify in order to improve student performance.

The findings of this paper carry implications for the potential future development of the transition countries' economies. The tested students have by now reached an age of about 21 and are about to enter the labor market. The measured schooling quality is thus to take effect on the countries' economies in the coming years. Besides the relatively advantageous economic and political situation, the first group of transition countries also presides over well functioning schooling systems. The high development of institutions is visible in the favorable incentive system in schools, where decentralization has widely progressed and positively affects student scores. Coming entrants into the labor markets are well educated, which should make them competitive to the EU labor market. The quick transformation of the schooling system in the early years of transition despite the financial and political hurdles is a good foundation for economic growth in the coming years. However, the high spread of test scores and the differentiation of students according to their background may result in a higher

inequality of income. A higher spread of income might be beneficial for short-term economic growth during the transition phase and might reflect entrepreneurial qualities of people, but it might threaten the social consensus in the long term.

The second group of transition countries instead has not fully reformed its schooling systems yet. The countries still feature many traits from communist times. Institutional reforms may need to be continued in the coming years. The transition will take longer than for the first group, and the labor markets may not yet be competitive to the markets of their Western neighbors due to the low quality and focus on homogeneity of the schooling systems. The low effect of student background on performance and the generally lower test-score variations speak against a rapid increase in the spread of the income distribution.

In the analysis of educational production in Eastern Europe, there is great scope for further research. It would be desirable to find further methods for bias control that leave more variation in test scores and are applicable to other inputs as well. Further, it would be interesting to study the development of other outcomes of education like social values conveyed to students, which are essential to form a steady democracy. Another path for further studies would be to compare the findings from TIMSS to other studies that were performed later, like TIMSS-Repeat in 1999 or PISA in 2000. This could reveal whether the trends of the schooling systems towards decentralization and choice have continued and whether the distinction between the two groups of countries still holds.

References

- Akerhielm, Karen (1995). Does Class Size Matter?. *Economics of Education Review* 14 (3): 229-241.
- Angrist, Joshua D. and Victor Lavy (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics* 114 (2): 533-575.
- Berryman, Sue (2000). *Hidden Challenges to the Education Systems in Transition Economies*. Washington, D.C.: The World Bank.
- Boeri, Tito and Katerine Terrell (2002). Institutional Determinants of Labor Reallocation in Transition. *Journal of Economic Perspectives* 16 (1): 51-76.
- Card, David and Alan B. Krueger (1996). Labor Market Effects of School Quality: Theory and Evidence. In: Gary Burtless (ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*: 43-73. Washington, D.C.: The Brookings Institution.
- Deaton, Angus (1997). *The Analysis of Household Surveys*. Published for the World Bank, Baltimore and London: The John Hopkins University Press.
- European Bank for Reconstruction and Development (EBRD) (2000). *Transition Report 2000 - Employment, Skills and Transition*. London.
- Figlio, David (1999). Functional Form and the Estimated Effects of School Resources. *Economics of Education Review* 18 (2): 241-252.
- Filer, Randall K. and Daniel Münich (2000). Responses of Private and Public Schools to Voucher Funding: The Czech and Hungarian Experience. CERGE-IE Working Paper No. 160.
- Gonzalez, Eugenio and Teresa Smith (eds.) (1997). *User Guide for the TIMSS International Database - Primary and Middle School Years*. International Association for the Evaluation of Educational Achievement, Chestnut Hill: TIMSS International Study Center, Boston College.
- Hanushek, Eric A. (2003). The Failure of Input-Based Schooling Policies. *Economic Journal* 113 (485): F64-F98.
- Hanushek, Eric A. and Javier Luque (2002). Efficiency and Equity in Schools Around the World. NBER Working Paper No. 8949.
- Hanushek, Eric A., Steven Rivkin and Lori Taylor (1996). Aggregation and the Estimated Effects of School Resources. *Review of Economics and Statistics* 78 (4): 611-627.
- Heyneman, Stephen (1997). Educational Choice in Eastern Europe and the Former Soviet Union: A Review Essay. *Education Economics* 5 (3): 333-339.
- Klazar, Stanislav, Milan Sedmíhradský and Alena Vančurová (2001). Returns of Education in the Czech Republic. *International Tax and Finance* 8 (4): 609-620.

- Krueger, Alan B. (2003). Economic Considerations and Class Size. *Economic Journal* 113 (485): F34-F63.
- Lazear, Edward (2001). Educational Production. *Quarterly Journal of Economics* 113 (3): 777-803.
- Mukherjee, Chandan, Howard White and Marc Wuyts (1998). *Econometrics and Data Analysis for Developing Countries*. London and New York: Routledge.
- Newell, Andrew and Barry Reilly (1999). Rates of Return to Educational Qualifications in the Transitional Economies. *Education Economics* 7 (1): 67-84.
- Schultz, Theodore W. (1975). The Value of the Ability to Deal with Disequilibria. *Journal of Economic Literature* 13 (3): 827-846.
- Summers, Anita A. and Barbara L. Wolfe (1975). Equality of Educational Opportunity Quantified: A Production Function Approach. Philadelphia Fed Research Papers, Department of Research, Federal Reserve Bank of Philadelphia.
- White, Halbert (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and Direct Test for Heteroskedasticity. *Econometrica* 48 (4): 817-838.
- Wooldridge, Jeffrey M. (2001). Asymptotic Properties of Weighted *M*-Estimators for Standard Stratified Samples. *Econometric Theory* 17 (2): 451-470.
- Wößmann, Ludger (2003). Schooling Resources, Educational Institutions, and Student Performance: The International Evidence. *Oxford Bulletin of Economics and Statistics* 65 (2): forthcoming.
- Wößmann, Ludger and Martin R. West (2002). Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS. Kiel Working Papers No.1099.

Table 1: Transitional Progress and Education Systems in the Eastern European Countries

	CSK	HUN	SVN	SLV	LTU	LVA	ROM
Population in '000 ^e	10,264	10,106	1,930	5,415	3,611	2,385	22,364
GDP/capita US\$ 1994 ^a	3,977	4,052	7,231	2,721	1,143	1,442	1,323
GDP/capita US\$ 2000 ^a	4,797	4,552	9,073	3,556	3,064	3,019	1,644
Estimated GDP level in 2000, 1989=100 ^b	98	104	114	103	65	64	77
Private sector share of GDP in %, 2001 ^b	80	80	65	80	70	65	65
No. of chapters closed provisionally ^c	19	22	21	19	18	16	8
Ed. exp. per student US\$ in 1994 ^d	671	840	1492	319	195	283	115
Student pop. in '000 ^f	1,146	1,360	189	804	512	299	2,461
Enrollment in 1994 ^g	99.5	99.1	96.7	97	92.2	89	91.4
No. of school types for lower secondary	2	2	1	4	5	1	1
Pre-primary education is partly obligatory	no	yes	yes	no	no	no	yes
Average attendance rate in pre-primary	86%	86.5%	59%	82%	40%	23.4%	65%

^a EBRD Transition Report 2001, Country Assessments.

^b EBRD Transition Report 2001, Table A3.1.

^c Acquis Communautaire chapters in negotiation on EU accession, EBRD Transition Report 2001, Box 2.2.

^d Total educational expenditure in 1994 US\$ divided by population of age 3 to 24 in 1994, UNESCO Statistical Yearbook 1997 and Berryman (2000).

^e Mid 2001 estimates, CIA Factbook Country Profiles.

^f No. of children of compulsory school age in 2000. See Eurybase.

^g Enrollment rates for basic education, ages 6/7 – 14/15 in 1994, Berryman (2000), Table A6.

Table 2: Student Background and Educational Performance

	CSK		HUN		SVN		SLV		LTU		LVA		ROM	
	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.
Upper Grade	70.92*	57.02*	63.49*	56.78*	63.05*	41.64*	61.06*	54.32*	67.25*	86.56*	48.51*	61.90*	33.90*	40.06*
	(5.92)	(5.54)	(3.09)	(3.44)	(4.90)	(4.18)	(4.19)	(4.08)	(4.44)	(4.56)	(4.24)	(4.25)	(3.93)	(5.11)
Age	-33.41*	-18.30*	-31.63*	-23.51*	-22.46*	-13.20*	-20.73*	-18.77*	-23.97*	-16.63*	-16.87*	-11.96*	-7.85*	-6.73 [†]
	(2.83)	(2.82)	(2.11)	(2.04)	(3.31)	(3.57)	(2.74)	(2.80)	(2.56)	(2.91)	(2.51)	(2.31)	(2.48)	(3.30)
Female	-13.34*	-25.70*	-7.00*	-21.47*	-7.10*	-22.13*	-5.99 [†]	-19.34*	4.36 [‡]	-9.72*	-6.92*	-14.20*	-2.55	-9.01*
	(3.25)	(2.72)	(2.36)	(2.28)	(2.42)	(2.64)	(2.39)	(2.45)	(2.67)	(2.64)	(2.41)	(2.42)	(2.55)	(2.92)
Immigrant	-22.07*	-14.45 [‡]	-7.26	-18.46 [‡]	1.87	-12.86*	-1.65	-11.64	-8.03	-14.02 [‡]	-14.29	-22.36 [†]	16.09 [†]	8.46
	(8.06)	(7.49)	(8.36)	(10.00)	(4.18)	(3.68)	(8.19)	(8.46)	(9.64)	(9.33)	(7.39)	(9.03)	(6.36)	(7.93)
Living with both parents	14.92*	9.39*	7.80 [‡]	9.06 [†]	2.36	-0.4	1.65	-4.42 [‡]	1.25	4.98	-7.1	-2.69	-3.52	-3.91
	(4.83)	(3.55)	(4.09)	(4.15)	(3.87)	(3.70)	(2.74)	(2.63)	(3.20)	(3.81)	(2.82)	(3.09)	(3.62)	(4.51)
<i>Parents' education</i>														
Finished Secondary	24.25*	14.82*	7.39 [‡]	6.54	23.99*	16.40*	15.43*	12.32*	11.02 [‡]	4.92	-4.7	2.43	1.93	-1.02
	(3.51)	(3.43)	(4.03)	(4.58)	(2.90)	(3.17)	(3.25)	(4.01)	(6.10)	(5.77)	(4.31)	(3.83)	(5.40)	(7.19)
Finished University	47.17*	33.59*	47.16*	37.18*	49.34*	36.08*	40.11*	35.56*	31.61*	23.10*	29.16*	25.13*	21.39*	20.85 [†]
	(4.68)	(4.67)	(6.14)	(5.76)	(4.22)	(4.71)	(4.51)	(4.68)	(6.59)	(6.55)	(5.32)	(4.52)	(7.03)	(9.02)
<i>Books at home</i>														
In ln	22.59*	20.17*	22.61*	19.98*	18.57*	17.42*	21.54*	19.14*	17.92*	13.36*	15.39*	10.05*	11.97*	12.68*
	(1.89)	(1.85)	(1.49)	(1.51)	(1.58)	(1.58)	(1.33)	(1.40)	(1.16)	(1.44)	(1.74)	(1.63)	(1.72)	(2.05)
<i>Community location</i>														
Close to center	5.75	1.81	8.90 [†]	10.29*	-2.78	-1.47	-1.8	-9.0	7.34 [‡]	2.71	4.69	1.00	19.98*	16.38 [†]
	(7.93)	(5.14)	(4.45)	(3.75)	(3.88)	(3.32)	(5.70)	(5.36)	(4.32)	(4.58)	(4.92)	(4.22)	(6.78)	(8.31)
cons	832.19*	671.24*	796.56*	724.60*	706.08*	633.38*	671.68*	673.21*	647.59*	556.97*	603.36*	544.20*	506.50*	494.08*
	(41.99)	(39.08)	(31.38)	(29.55)	(46.73)	(50.86)	(38.09)	(39.78)	(34.38)	(41.41)	(37.85)	(33.03)	(36.17)	(48.99)

Significance levels (based on robust standard errors): * 1 percent, [†] 5 percent, [‡]10 percent.

Separate least-squares regressions within each country and subject, weighted by students' sampling probabilities.

Dependent variable: TIMSS math/science test score. Clustering-robust standard errors in parentheses.

Table 3: Score Differences Worst/Best Student Background Scenario

	CSK	HUN	SVN	SLV	LTU	LVA	ROM
Math	219	183	152	151	126	113	112
Science	180	180	152	153	102	113	103

Table 4: Resources, Teacher Characteristics, and Educational Performance

	CSK		HUN		SVN		SLV		LTU		LVA		ROM	
	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.
Class size	2.75* (.88)	1.58† (.68)	.40 (.40)	.26 (.35)	.35 (.60)	-.47 (.56)	.12 (.50)	.84 (.72)	1.83* (.53)	-.25 (.45)	-.31 (.31)	-.43‡ (.25)	1.49* (.52)	.68 (.71)
Great short. of materials	-9.50‡ (5.02)	-7.08 (6.47)	6.54 (5.00)	6.65 (4.26)	-.06 (5.04)	-6.63 (4.41)	-7.29 (5.72)	2.09 (6.56)	-9.87‡ (5.17)	-8.93‡ (5.07)	5.04 (5.27)	9.07† (4.46)	-6.28 (9.02)	-10.51 (9.42)
<i>Teacher characterist.</i>														
Teacher is female	-5.02 (6.49)	-3.03 (6.52)	-4.72 (5.67)	2.72 (3.79)	-.03 (4.98)	1.32 (4.37)	4.06 (6.88)	6.85 (5.45)	5.12 (6.03)	6.75 (5.74)	-3.33 (8.17)	-1.68 (4.64)	8.65 (7.20)	3.35 (8.05)
Teacher's exp. in ln <i>Teacher's education</i>	1.34 (4.11)	-4.98 (3.05)	5.23† (2.29)	1.00 (2.25)	2.65 (3.36)	.78 (2.26)	7.51‡ (3.96)	1.46 (2.79)	2.03 (2.69)	.34 (2.31)	1.69 (2.55)	-1.20 (2.05)	8.91 (5.49)	9.78† (4.43)
BA or equiv.	No observ.	No observ.	-9.30 (6.52)	-2.36 (5.14)	3.69 (7.10)	9.46† (4.61)	No observ.	No variati on	4.82 (8.76)	-18.95‡ (10.04)	5.39 (10.51)	6.80 (4.57)	7.33 (6.44)	11.24 (7.24)
MA/Ph.D.	17.90 (13.93)	18.41† (7.92)	20.09 (15.00)	-29.50† (14.27)	No observ.	No observ.	-12.01‡ (6.21)	No variati on	5.84 (9.80)	-7.59 (10.64)	-16.44 (20.16)	-5.97 (14.03)	51.49* (9.13)	32.92* (12.48)

Significance levels (based on robust standard errors): * 1 percent, † 5 percent, ‡10 percent.

Separate least-squares regressions within each country and subject, weighted by students' sampling probabilities. Dependent variable: TIMSS math/science test score. Controlling for all student-background variables reported in Table 2. Clustering-robust standard errors in parentheses.

Table 5: Institutional Settings and Educational Performance

	CSK		HUN		SVN		SLV		LTU		LVA		ROM	
	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.	Math	Sci.
<i>School responsibilities</i>														
Autonomy	7.80 (4.97)	3.02 (3.40)	10.58 [‡] (5.76)	6.26 (4.02)	-3.09 (6.38)	-3.85 (5.26)	3.12 (3.25)	2.13 (3.01)	.77 (4.27)	-1.72 (4.77)	-6.43 (6.58)	-1.65 (4.73)	-3.45 (3.53)	-1.00 (4.51)
Determining teacher's salary	8.12 (6.95)	5.92 (5.36)	-5.72 (6.14)	-7.24 [‡] (4.11)	3.28 (4.25)	.65 (3.98)	-6.52 (6.89)	-3.66 (8.51)	-10.87 [†] (5.05)	-.37 (6.52)	-1.73 (5.41)	-7.79 (4.92)	-2.56 (10.65)	-6.23 (10.65)
<i>Strong infl. on curriculum</i>														
Teachers individually	14.32 [‡] (7.88)	3.62 (5.88)	7.36 (6.24)	4.86 (5.64)	-.46 (6.93)	4.75 (6.40)	-3.12 (9.26)	-2.11 (7.96)	6.23 (4.67)	6.44 (4.90)	-3.04 (4.79)	-.96 (4.13)	10.37 (11.24)	15.70 (10.47)
Teachers collectively	-7.92 (7.28)	-2.72 (5.08)	1.14 (5.52)	1.60 (5.08)	6.31 (8.91)	-1.10 (7.95)	2.86 (11.56)	.70 (9.94)	-1.59 (4.49)	-1.55 (5.20)	3.14 (5.20)	3.00 (4.33)	-6.53 (17.57)	-3.08 (19.86)
<i>Class teacher has strong influence on</i>														
Supplies or subject matter	-.90 (8.80)	-5.33 (4.39)	2.47 (4.55)	5.57 (3.48)	16.12 [*] (6.03)	4.28 (4.89)	-.51 (6.62)	-.32 (5.26)	14.43 (10.24)	-10.67 (6.65)	-1.12 (8.08)	-3.09 (5.15)	-8.03 (8.96)	-12.25 (9.75)
Kind supplies / textbooks	4.90 (5.17)	4.19 (4.70)	2.99 (5.35)	-6.67 [‡] (3.73)	-4.74 (3.76)	-5.07 (3.69)	7.33 (6.27)	-2.07 (4.99)	-2.72 (7.57)	6.77 (7.16)	1.33 (5.56)	-.49 (4.52)	19.41 [†] (8.28)	-11.69 (8.99)
<i>Additional</i>														
Homework	19.23 [‡] (9.86)	-6.62 (10.15)	3.22 (3.57)	-1.47 (2.68)	1.68 (2.02)	-4.82 (6.56)	2.07 (3.13)	9.26 (10.40)	-2.66 (2.12)	5.43 (3.41)	3.03 (2.75)	7.29 [†] (2.82)	1.93 (1.29)	10.07 [†] (4.18)
Unint. parents limit teaching	-19.79 [*] (6.62)	-19.11 [†] (9.47)	-8.40 (6.65)	1.84 (5.71)	.58 (5.19)	-7.43 [†] (3.71)	-1.38 (8.01)	-11.82 (7.45)	2.64 (6.63)	-7.69 (9.61)	-9.26 [‡] (4.99)	4.40 (5.62)	-.63 (5.05)	5.73 (6.29)

Significance levels (based on robust standard errors): * 1 percent, † 5 percent, ‡10 percent.

Separate least-squares regressions within each country and subject, weighted by students' sampling probabilities. Dependent variable: TIMSS math/science test score. Controlling for all student-background variables reported in Table 2 and for all resource variables reported in Table 4. Clustering-robust standard errors in parentheses.

Table 6: R² of Regressions with Different Categories of Explanatory Variables

	CSK	HUN	SVN	SLV	LTU	LVA	ROM
Math							
Student and family characteristics	0.1660	0.2121	0.1718	0.1424	0.2214	0.1208	0.0959
+ Resources	0.1776	0.2169	0.1723	0.1455	0.2324	0.1239	0.1158
+ Institutions	0.1996	0.2208	0.1770	0.1476	0.2379	0.1282	0.1253
Science							
Student and family characteristics	0.1466	0.1711	0.1270	0.1241	0.2386	0.1356	0.0808
+ Resources	0.1546	0.1730	0.1296	0.1264	0.2460	0.1404	0.0949
+ Institutions	0.1615	0.1761	0.1321	0.1284	0.2505	0.1443	0.1111

Table 7: Class-Size Coefficients for Class Segments

	CSK	HUN	SVN	SLV	LTU	LVA	ROM	EAST
All students								
Math class	2.23 [†] (.94)	.48 (.39)	.44 (.60)	.22 (.51)	1.64* (.53)	-.36 (.32)	1.38* (.48)	.80* (.20)
Science class	1.25 [‡] (.69)	.25 (.35)	-.51 (.55)	.86 (.67)	-.19 (.44)	-.43 [‡] (.26)	.69 (.69)	.37 [‡] (.20)
Lower segment								
Math class	.30 (1.37)	-.25 (.93)	-.32 (1.82)	.03 (1.13)	1.50 (1.09)	-1.35 (.83)	-.96 (1.19)	.24 (.40)
Science class	-.51 (.93)	-.17 (.83)	-3.07 [†] (1.23)	.85 (1.02)	1.46 (1.23)	-1.56 (.97)	-1.39 (1.49)	-.46 (.41)
Middle segment								
Math class	16.77* (5.00)	5.18 [‡] (2.61)	9.93* (3.72)	-6.53 (4.48)	-.43 (7.28)	-1.96 (2.19)	-.80 (2.91)	3.78* (1.42)
Science class	-4.92 (5.91)	-3.32 (2.63)	7.95 [‡] (4.31)	-5.34 (3.64)	4.42 (6.45)	-2.35 (4.03)	-1.34 (3.10)	2.65 (1.79)
Upper segment								
Math class	.24 (2.91)	2.69 (2.04)	3.44 (2.10)	1.47 (3.23)	4.36* (1.50)	-2.07 [†] (.78)	4.73* (1.45)	1.55* (.64)
Science class	4.70 [†] (2.16)	2.86 [†] (1.27)	-.87 (1.68)	1.28 (2.48)	-1.27* (.45)	-1.13* (.31)	5.19* (1.79)	.57 (.39)
Mean Math class	25.35	22.41	24.67	26.09	20.86	21.65	26.67	23.95
Mean Science class	25.52	22.14	24.42	26.54	21.63	23.74	26.23	24.32

Significance levels (based on robust standard errors): * 1 percent, [†] 5 percent, [‡] 10 percent.

Separate least-squares regressions within each country and subject, weighted by students' sampling probabilities. Dependent variable: TIMSS math/science test score. Controlling for all student-background variables reported in Table 2, for all other resource variables reported in Table 4, and for all institutional variables reported in Table 5. Clustering-robust standard errors in parentheses.

Table 8: The Coefficient on Class Size in Different Models

	CSK	HUN	SVN	SLV	LTU	LVA	ROM
Math class							
Least squares	2.23 [†] (.94)	.48 (.39)	.44 (.60)	.22 (.51)	1.64* (.53)	-.36 (.32)	1.38* (.48)
School fixed effects	-.98 (.62)	.06 (.46)	-.08 (.63)	.06 (.43)	1.69* (.64)	-1.00* (.29)	1.65* (.33)
SFE + IV	.89 (1.63)		.74 (1.19)				-.48 (4.19)
Science class							
Least squares	1.25 [‡] (.69)	.25 (.35)	-.51 (.55)	.86 (.67)	-.19 (.44)	-.43 [‡] (.26)	.69 (.69)
School fixed effects	.05 (.61)	.04 (.48)	-.50 (.64)	.84 (.52)	-.12 (.38)	.29 (.25)	.65 [‡] (.37)
SFE + IV	-.38 (1.29)		-.20 (1.14)				-.15 (1.79)
Students LS / SFE	6672/6659	5978/5962	5606/5576	7101	5056	4976/4917	7471/7462
Schools LS / SFE	150/149	150/149	122/121	145	145	143/141	163/162

Significance levels (based on robust standard errors): * 1 percent, [†] 5 percent, [‡] 10 percent.

Separate regressions within each country and subject, weighted by students' sampling probabilities. Dependent variable: TIMSS math/science test score. Controlling for all student-background variables reported in Table 2. Clustering-robust standard errors in parentheses.

Table 9: Educational Production in Eastern and Western Europe

	East	First	Second	West		East	First	Second	West
Student and family characteristics					Institutional setting				
Upper Grade	58.36*	66.05*	49.78*	55.72*	<i>School responsibilities</i>				
	(1.71)	(2.31)	(2.49)	(2.14)	Autonomy	1.71	5.25 [†]	-3.06	3.17
Age	-22.52*	-28.25*	-16.22*	-20.75*	(budget, suppl., t.)	(1.85)	(2.26)	(2.75)	(3.65)
	(1.03)	(1.39)	(1.49)	(1.31)	Determining	1.35	3.79	-5.00	-5.21
Female	-5.22*	-8.07*	-1.83	-9.89*	teacher salary	(2.55)	(3.21)	(4.03)	(4.53)
	(1.00)	(1.34)	(1.49)	(1.47)	<i>Strong influence on curriculum</i>				
Immigrant	5.71 [‡]	-2.46	10.15 [†]	-25.63*	Teachers	4.48 [‡]	5.17	3.75	3.77
	(3.16)	(3.00)	(4.66)	(3.20)	individually	(2.51)	(3.82)	(3.27)	(4.20)
Living with both parents	.75	6.29*	-1.08	5.44*	Teachers	-1.61	-2.26	-.07	-14.05*
	(1.45)	(1.91)	(1.91)	(1.57)	collectively	(3.01)	(4.41)	(3.65)	(3.57)
Parents' education					<i>Math class teacher has strong influence</i>				
Finished Secondary	11.07*	17.79*	-.25	9.10*	Money for suppl. or subject matter	2.55	3.72	-.59	-.33
	(1.76)	(1.79)	(3.30)	(2.23)	Kind of supplies or textbooks	(2.83)	(3.37)	(5.28)	(2.82)
Finished University	37.80*	45.04*	24.15*	28.91*	Homework	3.51	2.60	6.93 [‡]	.98
	(2.10)	(2.44)	(3.73)	(2.90)	Uninterested parents limit teaching	(2.27)	(2.70)	(3.96)	(2.50)
Books at home	16.68*	20.72*	13.62*	14.93*	Intercept	2.25 [†]	3.71 [‡]	1.05	8.18*
	(.67)	(.78)	(.95)	(.68)		(.99)	(1.95)	(1.09)	(1.85)
Close to the center	3.59	2.64	6.81 [†]	3.84		-4.09 [‡]	-5.02	-2.72	-21.00*
	(2.20)	(2.86)	(3.24)	(3.33)		(2.43)	(3.55)	(3.22)	(5.51)
Resources and teacher characteristics						598.19*	706.45*	546.48	664.74*
Math Class size	.80*	.68 [†]	.81*	1.62*	Students	(18.26)	(24.20)	(26.57)	(23.70)
	(.20)	(.28)	(.27)	(.37)	Schools				
Great shortage of materials	-2.19	.58	-3.11	3.85		42,815	25,357	17,458	21,933
	(2.55)	(3.17)	(3.60)	(3.45)					
<i>Math Teacher char. and education</i>						1,017	567	450	553
Teacher is female	.69	-2.07	5.36	-4.83 [‡]					
	(2.73)	(3.24)	(4.50)	(2.69)					
Teacher's exper. in ln	5.04*	4.21 [†]	5.43 [†]	2.76					
	(1.47)	(1.95)	(2.26)	(2.04)					
BA or equivalent	1.59	-6.86	5.12	-3.47	Mean math score	500.97	527.09	466.14	505.10
	(3.97)	(5.39)	(4.94)	(4.81)		(93.61)	(91.93)	(84.00)	(86.34)
MA/Ph.D.	4.37	14.46	4.97	-4.91					
	(6.00)	(9.11)	(7.68)	(5.86)					

Significance levels (based on robust standard errors): * 1 percent, [†] 5 percent, [‡]10 percent.

Separate least-squares regressions within each country group, weighted by students' sampling probabilities.

Dependent variable: TIMSS math test score. Clustering-robust standard errors in parentheses.

Appendix

Table A1 Participation of Students, Classes and Schools in TIMSS

	CSK	HUN	SVN	SLV	LTU	LVA	ROM
Students	6672	5978	5606	7101	5056	4976	7471
Classes	299	299	243	290	292	284	325
Schools	150	150	122	145	145	143	163
Sampled students (in percent) ^a	0.6%	0.4%	3.0%	0.9%	1.0%	1.7%	0.3%

^a Number of sampled students over the number of children of compulsory school age in percent.

Table A2 Definition of Variables and Range of Values

Variable Name	Definition	Type	Min	Max
Math Score	International Math Test Score	Numeric	141.6	887.44
Science Score	International Science Test Score	Numeric	85.02	872.46
Student and family characteristics				
Upper Grade	Grade level of students	Dummy	0 for 7 th grade	1 for 8 th grade
Age	Age of students	Numeric	10.1	20.4
Female	Sex of students	Dummy	0 for male	1 for female
Immigrant	Origin of students	Dummy	0 for other students	1 for immigrated student or parent
Living with both parents	Student's family situation	Dummy	0 for one or no parent	1 for both parents
Parents' education				
No or some secondary		Dummy	Ref.	Ref.
Finished Secondary	Highest educational level reached by a parent	Dummy	0 for ref.	1 for finished sec. and some after
Finished University		Dummy	0 for ref.	1 for finished univ.
Books at home				
In ln	Number of books at student's home in ln	categorical	1.61	5.52
Community location				
Non-urban area		Dummy	Ref.	Ref.
Close to the center of a town	Location of the student's community	Dummy	0 for ref.	1 for close to center
Resources and teacher characteristics				
Class size	Size of student's class	Numeric	3	58
No or some shortage of materials		Dummy	Ref.	Ref.
Great shortage of materials	Degree of school's shortage of materials	Dummy	0 for ref.	1 for great shortage
<i>Teacher characteristics</i>				
Teacher is female	Class teacher's sex	Dummy	0 for male	1 for female
Teacher's exp. in ln	Years of class teacher's experience in ln	Numeric	0	3.91
<i>Teacher's education</i>				
Secondary and less		Dummy	Ref.	Ref.
BA or equivalent	Highest educational level of class teacher	Dummy	0 for ref.	1 for Bachelor/equ.
MA/Ph.D.		Dummy	0 for ref.	1 for Master/equ.

Table A2 Definition of Variables and Range of Values (continued)

Variable Name	Definition	Type	Min.	Max.
Institutional setting				
<i>School responsibilities</i>				
Autonomy (budget, supplies, teachers)	Degree of school's autonomy	Categorical	0 for autonomy in no field	3 for autonomy in all three fields
Determining teacher salary	School's responsibility over setting salaries	Dummy	0 for no resp.	1 for resp. over salary
<i>Strong influence on curriculum</i>				
Teachers individually	Influence of individual or subject teachers on curr.	Dummy	0 for no infl.	1 for infl. of at least one group
Teachers collectively	Infl. of teachers collect. or teacher unions on curr.	Dummy	0 for no infl.	1 for infl. of at least one group
<i>Class teacher has strong influence on</i>				
Money for supplies or subject matter	Strong infl. of class teachers	Dummy	0 for no strong infl.	1 for strong infl. on at least one field
Kind of supplies or textbooks	Strong infl. of class teachers	Dummy	0 for no strong infl.	1 for strong infl. on at least one field
<i>Additional</i>				
Homework	Homework for subject in hours per week	Numeric	1	9.6
Uninterested parents limit teaching	Teacher notices uncooperative parents negatively	Dummy	0 for no negative infl. on students	1 for negative infl. on students

Min. and max. values include all seven Eastern European countries.

Table A3 Means and Standard Deviations of the Data

	CSK	HUN	SVN	SLV	LTU	LVA	ROM
Math Score	543.57 (93.70)	519.05 (93.52)	518.68 (87.43)	527.08 (90.73)	453.72 (81.60)	476.68 (80.82)	468.01 (87.81)
Science Score	553.44 (86.90)	535.30 (92.48)	544.42 (88.11)	526.69 (90.16)	441.27 (88.26)	458.59 (83.35)	468.85 (102.33)
Student and family characteristics							
Upper Grade	.50 (.50)	.49 (.50)	.48 (.50)	.49 (.50)	.52 (.50)	.47 (.50)	.50 (.50)
Age	13.89 (.65)	13.82 (.71)	14.27 (.63)	13.77 (.62)	13.84 (.66)	13.77 (.72)	14.12 (.71)
Female	.50 (.50)	.50 (.50)	.51 (.50)	.51 (.50)	.52 (.50)	.52 (.50)	.51 (.50)
Immigrant	.02 (.14)	.02 (.16)	.11 (.32)	.02 (.14)	.02 (.15)	.02 (.14)	.13 (.34)
Living with both parents	.89 (.32)	.92 (.27)	.90 (.30)	.83 (.37)	.85 (.36)	.80 (.40)	.74 (.44)
Parents' education							
Finished Secondary	.55 (.50)	.72 (.45)	.64 (.48)	.58 (.49)	.56 (.50)	.61 (.49)	.57 (.50)
Finished University	.21 (.41)	.19 (.39)	.19 (.39)	.22 (.41)	.37 (.48)	.29 (.45)	.09 (.29)
Books at home							
In ln	4.79 (.73)	4.73 (.99)	4.34 (.97)	4.41 (.87)	4.31 (1.04)	5.03 (.75)	3.66 (1.46)
Community location							
Close to the center of a town	.40 (.49)	.45 (.50)	.38 (.48)	.29 (.45)	.61 (.49)	.30 (.46)	.51 (.50)
Resources and teacher characteristics							
Math Class size	25.35 (3.36)	22.41 (5.19)	24.66 (3.70)	26.09 (4.32)	20.86 (3.93)	21.65 (6.75)	26.67 (6.46)
Science Class size	25.52 (3.41)	22.14 (4.98)	24.42 (3.47)	26.54 (3.48)	21.63 (5.39)	23.75 (7.80)	26.23 (6.49)
Great shortage of materials	.03 (.18)	.17 (.38)	.20 (.40)	.02 (.14)	.22 (.42)	.63 (.48)	.17 (.37)
<i>Math Teacher characteristics</i>							
Teacher is female	.84 (.36)	.85 (.36)	.88 (.32)	.79 (.41)	.85 (.35)	.94 (.25)	.67 (.47)
Teacher's exp. in ln	2.82 (.83)	2.69 (.75)	2.63 (.56)	2.86 (.64)	2.79 (.70)	2.65 (.73)	2.84 (.74)
<i>Math Teacher's education</i>							
BA or equivalent	0 (0)	.09 (.29)	.05 (.22)	0 (0)	.81 (.39)	.93 (.25)	.45 (.50)
MA/Ph.D.	.99 (.11)	.02 (.15)	0 (0)	.99 (.08)	.16 (.36)	.02 (.14)	.00 (.07)

Table A3 Means and Standard Deviations of the Data (continued)

	CSK	HUN	SVN	SLV	LTU	LVA	ROM
<i>Science Teacher characteristics</i>							
Teacher is female	.77 (.42)	.78 (.42)	.82 (.38)	.66 (.47)	.85 (.36)	.83 (.38)	.77 (.42)
Teacher's exp. in ln	2.75 (.95)	2.69 (.81)	2.63 (.70)	2.69 (.81)	2.72 (.84)	2.61 (.88)	2.77 (.81)
<i>Science Teacher's education</i>							
BA or equivalent	0 (0)	.21 (.41)	.14 (.34)	0 (0)	.72 (.45)	.88 (.32)	.54 (.50)
MA/Ph.D.	.96 (.19)	.01 (.08)	0 (0)	1 (0)	.22 (.42)	.03 (.16)	.01 (.08)
Institutional setting							
<i>School responsibilities</i>							
Autonomy (budget, supplies, teachers)	2.59 (.66)	2.92 (.30)	2.98 (.13)	2.40 (.80)	2.79 (.44)	2.85 (.39)	1.51 (.86)
Determining teacher salary	.59 (.49)	.88 (.33)	.29 (.45)	.85 (.36)	.15 (.36)	.75 (.44)	.07 (.25)
<i>Strong influence on curriculum</i>							
Teachers individually	.83 (.37)	.79 (.41)	.24 (.42)	.16 (.37)	.50 (.50)	.44 (.50)	.08 (.27)
Teachers collectively	.66 (.47)	.54 (.50)	.15 (.36)	.10 (.30)	.24 (.43)	.56 (.50)	.04 (.20)
<i>Math class teacher has strong influence on</i>							
Money for supplies or subject matter	.13 (.33)	.77 (.42)	.12 (.33)	.19 (.39)	.04 (.20)	.12 (.33)	.17 (.37)
Kind of supplies or textbooks	.42 (.49)	.22 (.41)	.34 (.47)	.16 (.37)	.07 (.26)	.23 (.42)	.18 (.38)
<i>Science teacher has strong influence on</i>							
Money for supplies or subject matter	.16 (.37)	.71 (.46)	.17 (.38)	.27 (.44)	.14 (.35)	.26 (.44)	.21 (.40)
Kind of supplies or textbooks	.50 (.50)	.29 (.45)	.50 (.50)	.22 (.41)	.13 (.34)	.26 (.44)	.17 (.37)
<i>Additional Math</i>							
Homework	0.58 (0.43)	1.36 (.51)	1.89 (.93)	1.18 (.71)	2.06 (.90)	1.87 (.74)	4.52 (1.83)
Uninterested parents limit teaching	.07 (.26)	.08 (.27)	.18 (.38)	.11 (.31)	.09 (.28)	.17 (.38)	.50 (.50)
<i>Additional Science</i>							
Homework	.16 (.18)	.77 (.65)	.28 (.25)	.27 (.20)	.63 (.58)	.59 (.53)	.76 (.95)
Uninterested parents limit teaching	.06 (.23)	.10 (.30)	.17 (.38)	.08 (.28)	.06 (.24)	.09 (.28)	.37 (.48)

Standard deviations are reported in parentheses.

Values are weighted by the sampling probability of the students.