



OESTERREICHISCHE NATIONALBANK

Stability and Security.

WORKSHOPS

Proceedings of OeNB Workshops

*Strategies
for Employment and Growth
in Austria*

March 3, 2006



No. 10

Growth, Human Capital and the Quality of Schools: Lessons from International Empirical Research

Ludger Wößmann¹

University of Munich and Ifo Institute for Economic Research

1. Economic Growth and the Quality of Schooling

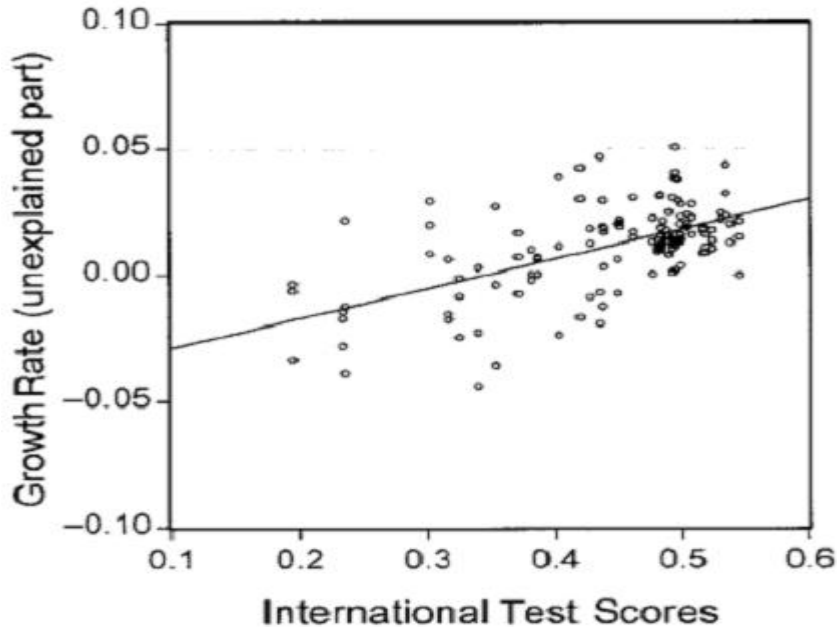
Human capital in the form of education is one of the driving forces in the long-run economic growth of countries. Importantly, it is less the mere *quantitative* educational attainment in terms of average years of schooling that drives economic performance – although this is certainly of importance, as demonstrated by de la Fuente (2004). What is even more important is the *quality* of schooling, as measured by performance on cognitive achievement tests, which has been shown to exert an even stronger impact on long-run economic growth and the level of economic development (Hanushek and Kimko, 2000; Barro, 2001; Wößmann, 2003d; Hanushek, 2005).

Chart 1, taken from Barro (2001), depicts the significantly positive effect of international test scores of student achievement on growth rates of real gross domestic product (GDP) per capita between 1965 and 1995 (for all countries with international test-score data), after other effects such as those of the initial level of GDP, government consumption, the rule of law, international openness, fertility, investment and others have been controlled for. The results reveal that, while both the quantity and the quality of schooling matter for economic growth, quality is much more important. Similarly, Wößmann (2003d), building on Gundlach et al. (2002), finds that once the quality of schooling in terms of test-score performance is taken into account, the share of cross-country variation in levels of economic development, measured by output per worker in 132 countries in 1990, that can be attributed to international differences in human capital rises from 21% to 45%

¹ University of Munich and Ifo Institute for Economic Research, Poschingerstr. 5, 81679 Munich, Germany. Tel.: ++49-89/9224-1699, Fax: ++49-89/9224-2699, E-mail: woessmann@ifo.de, Internet: www.cesifo.de/link/woessmann-l-e.htm.

(chart 2), and to over 60% in samples with reasonable data quality. Thus, the quality of schooling seems to be a crucial part of the human capital of a country.

Chart 1: Student Achievement and Economic Growth



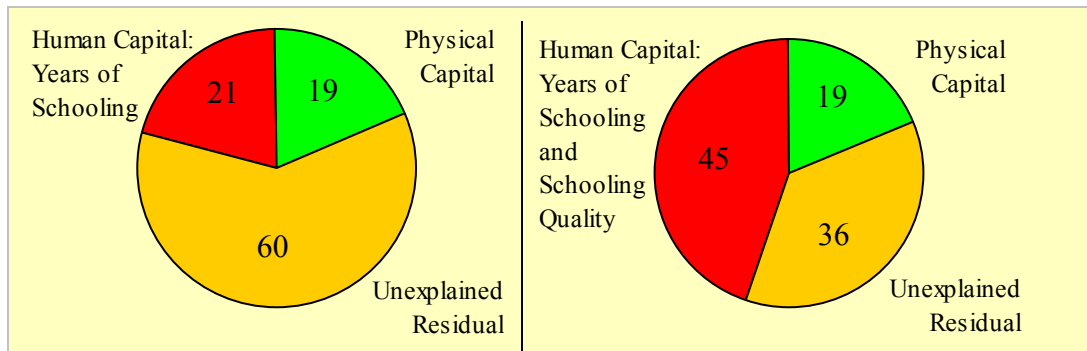
Source: Barro (2001).

How, then, can the quality of schooling be positively affected? Educational administrators and policymakers often argue that more *resources* would be needed for students to acquire more competencies. However, ample evidence shows that just increasing spending within current education systems is unlikely to improve students' performance substantially. Overwhelming evidence shows that expansions on the input side, such as simple physical expansion of the educational facilities and increased spending per student, generally do not seem to lead to substantial increases in children's competencies and learning achievement.² The same pattern also holds across countries: Students in countries with higher spending levels or smaller classes do not tend to perform better than students in less well equipped countries (cf. Wößmann, 2002, 2003a; Fuchs and Wößmann, 2004b, 2006). Even the equipment with computers in the classroom is not

² For evidence on the lack of substantial resource effects in general, and class-size effects in particular, cf., e.g., Gundlach et al. (2001), Hanushek et al. (1994), Hanushek (2003), Wößmann (2002; 2005c) and Wößmann and West (2006).

significantly associated with students' learning achievement (Fuchs and Wößmann, 2004a).³

Chart 2: Decomposition of International Differences in Output per Worker



Source: Based on results in Wößmann (2003d).

The lack of resource effects leaves open the question how we can improve the quality of schooling. Are there more constructive policy conclusions? Economic theory suggests that the performance of a system is affected by the *incentives* that actors face. That is, if the actors in the education process are rewarded for producing better student performance, and if they are penalized for not producing high performance, this will improve performance. The incentives to produce high-quality education, in turn, are created by the *institutions* of the education system – all the rules and regulations that set rewards and penalties (or fail to do so) for the people involved in the education process. Therefore, we might expect that institutional features have important impacts on student learning.

Recent research shows that such institutional features are indeed very important, suggesting that institutional reforms of the education system itself seem to be required to face the challenge of providing high-quality education effectively. Three institutional features that may be part of a successful system providing students with capabilities are the competition introduced by private-sector participation, decentralization of responsibilities that gives autonomy to schools and features such as centralized exams that make schools accountable to citizens and administrators (cf. Wößmann, 2004). If rightly pursued, all these institutional reforms can focus attention on learning achievement by directing stakeholders' incentives towards creating competencies for students.

³ By contrast, all studies on international educational performance find strong family-background effects on educational performance, with students from better-educated homes with a higher socio-economic status performing substantially better (cf., e.g., Wößmann, 2003a; Fuchs and Wößmann, 2004b, 2006; Schütz et al., 2005). Unfortunately, these family-background features are not subject to easy policy control.

However, *evidence* on the effects of such institutional features is hard to come by, particularly because systemic features such as competition, autonomy and accountability usually do not vary much within individual countries. For example, central exams, which are one mechanism to introduce accountability, tend to be a national feature, so that they are either present in the whole country or not at all.⁴ Furthermore, choice and accountability can often be expected to exert their impact in a systemic way, affecting not only individual schools but the whole system. For example, the prevalence of private schools may not only affect the performance of students in these private schools, but also the performance in public schools that are located nearby and exposed to the competition of the private schools. Take the Dutch school system as an example. The fact that three quarters of Dutch students attend privately managed schools may exert systemic effects for the whole Dutch school system, relative to school systems with small shares of private schools. Another problem with evidence from within individual countries is that where such within-country institutional variation exists, it is often not random but purposefully introduced by choices of individuals who may also differ along other lines, thereby confounding any empirical identification of the actual effects of the institutional features.⁵

Therefore, the research reported in this paper looks at a different kind of variation in the prevalence of competition, autonomy and accountability: The variation that exists *across* countries. For example, it asks whether students perform better in terms of their educational knowledge in countries where parents have a lot of choice to send their children to privately managed schools. To answer this kind of questions, the paper uses data from several recent international student achievement tests, which provide information on students' educational achievement that is comparable across many countries. Thereby, the research jointly looks at as many countries as possible, in order to analyze what countries can learn from each other in terms of the effects of competition, autonomy and accountability.

In section 2, the paper briefly sketches the theoretical argument why institutions should matter for the educational achievement of students, argues in favor of international variations to estimate the effects of institutions and briefly describes the four international student achievement tests that provide the data for the analyses discussed. Section 3 then discusses the evidence on the effects of different institutional features on the quality of schools, as measured by students' educational performance. It starts with evidence on competition from private schools. Next, it looks at decentralization of the education system, including the

⁴ Exceptions are Canada and Germany, where central exams are a regional feature.

⁵ Recent examples of studies based on the kind of variation in competition, autonomy and accountability that exists within countries, and which attempt to make sure that the estimates are not confounded by other effects, will be discussed in the appropriate sections below.

effects of devolving authority away from central authorities to local providers and of participation of parents and local communities in the supervision of schools. Finally, it discusses evidence on the effects of making schools accountable, ending with evidence on the complementarity between external exams (as accountability devices) and school autonomy (in decentralized school systems). Section 4 sums up the lessons from international empirical research on how to improve the quality of schools, which – as an advancement of a country’s human capital – could foster long-run economic growth.

2. The Quality of Schools and Institutions of the Education System

2.1 Why Should Institutions Matter?

Why would we expect, from a theoretical point of view, that institutions that introduce competition, autonomy and accountability might have an effect on student learning?⁶ The background of these considerations is that in the private business sector, market competition tends to discipline firms to work effectively because they would otherwise fail to profit. Inefficiency leads to higher costs and higher prices – practically an invitation to competitors to lure away customers.

However, all over the world, countries finance and manage the great majority of their schools publicly (cf. Pritchett, 2002). This relative lack of competition in the compulsory education sector tends to dull incentives to improve quality and restrain costs (cf. Hanushek et al., 1994). Moreover, in the public system, the ability of parents and students to ensure that they receive a high-quality education is often constrained by enormous obstacles to leaving bad schools.

This is the reason why institutions that ensure choice between autonomous schools and accountability of these schools may be expected to improve school quality in terms of student performance. Such institutions create incentives for school personnel to use their resources in ways that maximize performance, so that they may ultimately improve student learning.⁷

The choice and accountability that different institutions can introduce is not limited to the choice for parents in terms of the availability of privately managed schools. It also includes, for example, choice for schools and teachers in terms of their ability to make autonomous decisions. Likewise, accountability may be aimed at schools or at students, through such institutional features as external exit examinations and regular monitoring of student progress by tests and exams.

⁶ Sections 2 and 3 draw from Wößmann (2005b) in many parts.

⁷ Cf. Bishop and Wößmann (2004) for a more elaborate theoretical model of institutional effects in education.

2.2 How to Get Evidence on the Effects of Institutions?

How can we test whether these hypothesized effects of competition, autonomy and accountability prevail in the real world? And how can we estimate how large the effects are? To get evidence on the institutional effects, one needs *variation* in the institutional factors. For example, you want to compare whether somebody who has choice performs differently from somebody who does not have choice. Lacking such variation, one obviously cannot provide evidence on the effects: Comparing two persons who both have choice, or two persons who both do not have choice, cannot answer whether choice (or the lack thereof) had an effect on their performance.

Variation in institutional factors such as competition, autonomy and accountability are often not given *within a single country*: You either have it or you don't. This is most apparent in the case of system-wide central exams, which are either given for all students in the system or for none. If so, there is no way to provide evidence on the performance effect of this institution from within a country, because one can only compare persons who are all "treated" by central exams or only persons who are all not "treated". Because most of the existent research tends to focus on individual countries, the potentially important effects of choice and accountability tend to be missed in most empirical studies of the determinants of educational performance.

So, how can we then get evidence on institutional effects? The road taken in this paper is: There is institutional variation *across countries*. Some countries have central exam systems, others not. People in some countries are free to choose their schools, while people in other countries are not. This paper uses this kind of variation to see which institutional factors are related to better student learning, and which not. For example, it estimates whether students show better educational performance in countries where parents and schools have a certain kind of choice relative to students in countries where parents and schools do not have this kind of choice.

2.3 The Data: International Student Achievement Tests

The data that enable this cross-country identification of institutional effects are international student achievement tests. These tests quantify the educational performance of students in subjects such as math, science and reading by using the same test items in all participating countries. Thus, they provide measures of educational performance which are directly comparable across countries. Furthermore, by using representative sampling methods to draw random samples of schools, all the international student achievement tests used in this paper provide representative samples of students in each participating country.

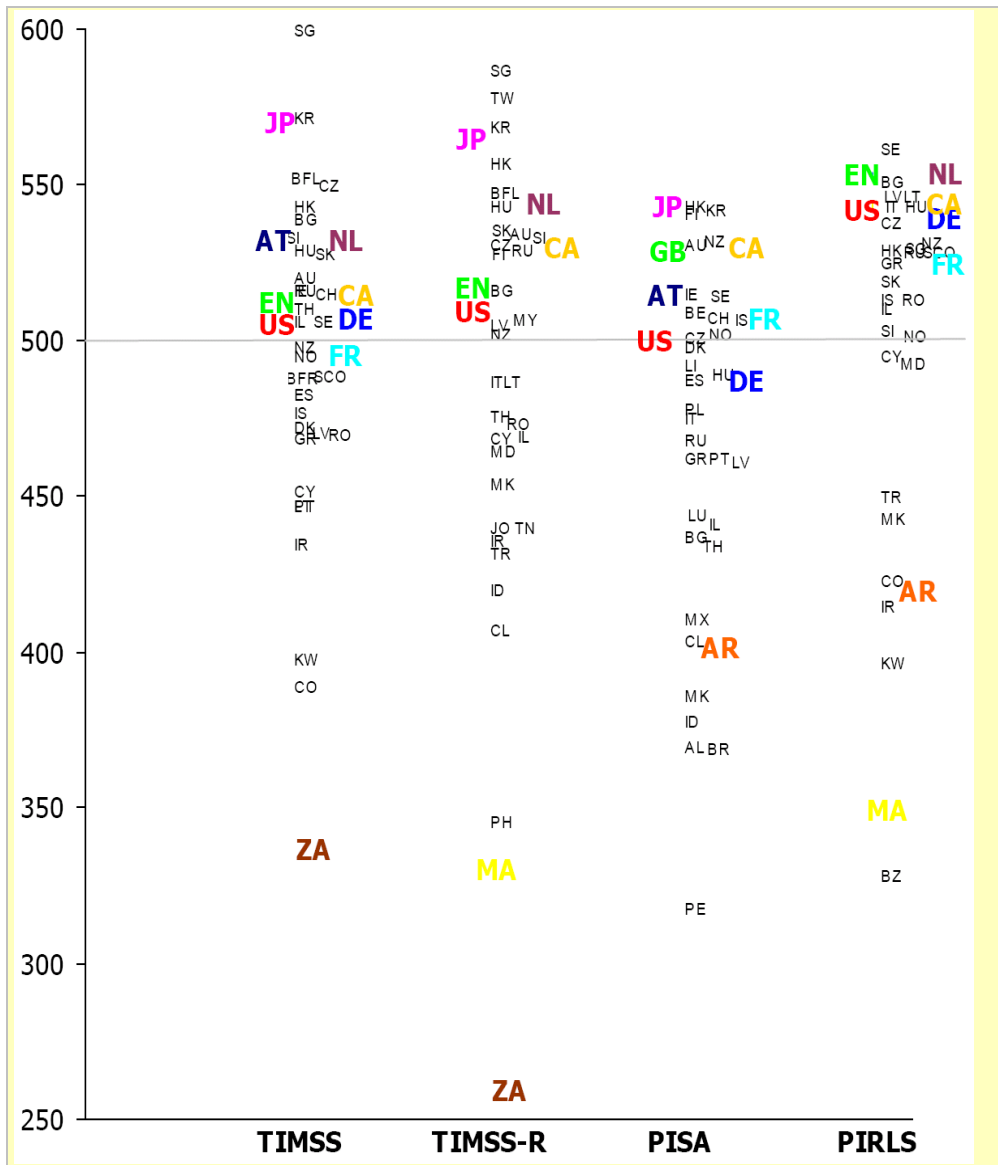
In particular, the research summarized in this paper uses data from four different recent international student achievement tests. The first one is the Third International Mathematics and Science Study (TIMSS), conducted in 1995 with data released in 1997. TIMSS was conducted by the International Association for the Evaluation of Educational Achievement (IEA), an independent cooperation of national research institutes and governmental research agencies. TIMSS targeted representative samples of students in the two adjacent grades with the largest share of 13-year-olds (usually 7th and 8th grade). For the analyses conducted in this paper, TIMSS yielded internationally comparable data for 266,545 students from 6,107 schools in 39 countries (for details, cf. Wößmann, 2003a and the references therein). Second, the IEA replicated the TIMSS test in 1999 under the name TIMSS-Repeat, with data released in 2001. TIMSS-Repeat targeted the upper of the two grades tested in TIMSS (usually the 8th grade), covering 180,544 students in 38 countries (cf. Wößmann, 2003b and the references therein). The sample of participating countries differed considerably between the two tests, so that the pooled TIMSS/TIMSS-Repeat database contains 54 different countries (447,089 students).

Third, the Organisation for Economic Co-operation and Development (OECD) conducted the Programme for International Student Assessment (PISA) in 2000, with data released in 2002, which targeted fifteen-year-old students. The PISA database covers 175,227 students in reading (96,855 in math, 96,758 in science) in 32 countries (cf. Fuchs and Wößmann, 2006 and the references therein). Fourth, in 2001 the IEA conducted the Progress in International Reading Literacy Study (PIRLS), with data released in 2003. While the focus of the previous studies was on secondary schools, PIRLS tested the reading performance of 140,626 primary-school students in 35 countries (cf. Fuchs and Wößmann, 2004b and the references therein). The target population of PIRLS was the upper of the two grades with the highest share of 9-year-olds of a country (usually the 4th grade).

Chart 3 provides a plot of the aggregate performance of the countries participating in each of the four tests. Each test was scaled so as to yield an international mean performance of 500 among the countries participating in the respective test, with an international standard deviation of 100.⁸ As is evident from chart 3, Austria performed on a quite respectable level in TIMSS and PISA, although the results in the 2003 cycle of PISA were significantly lower than in the 2000 cycle depicted in chart 3.

⁸ In PISA, the mean of 500 was scaled for the group of OECD countries only. As a consequence, the mean of all countries participating in PISA is somewhat lower than 500.

Chart 3: Aggregate Performance on International Student Achievement Tests



Note: The two-letter acronyms are the ISO codes of participating countries as coded by the International Organization for Standardization. Examples: AT = Austria; AR = Argentina; CA = Canada; DE = Germany; EN = England; FR = France; GB = Great Britain; JP = Japan; MA = Morocco; NL = Netherlands; US = United States; ZA = South Africa.

Source: Author's depiction based on data from the four tests.

The question addressed in this paper is whether, on average, the countries performing better than Austria on these tests feature an institutional set-up of their education systems that gives a bigger role to competition, autonomy and accountability, after holding constant other influence factors such as parental background, the development level of a country and the mean educational expenditure per student of a country. Given that the Netherlands (the country with the largest share of privately managed schools) and Japan (the country with the largest share of private schools that are also financially independent from government funding)⁹ are two countries that perform consistently better than the mean, there seems to be some preliminary indication that choice might matter for student performance.

However, the research presented in this paper goes far beyond comparing the aggregate performance across countries. Rather, it analyzes performance at the level of the individual student, using individual student-level data not only on educational performance in math, science and reading, but also combining it with extensive background information on other potential influence factors. These include dozens of indicators of family background, mostly taken from student background questionnaires (and parental background questionnaires in the case of primary-school PIRLS); several indicators of the resource endowment of the specific class or school, mostly taken from teacher and school background questionnaires; and several indicators of institutional features of the school systems, mostly taken from school background questionnaires. Among the latter are several indicators of the extent of competition, autonomy and accountability in the specific school of each tested student.

3. International Evidence for Institutional Effects on Schooling Quality

To estimate the effects of institutions that introduce competition, autonomy and accountability, the research summarized in this paper employs econometric techniques that control for differences in family background and the level of resources devoted to education.¹⁰ What do these studies of international achievement tests find out in terms of the effects of the different institutions introduced above on the quality of schooling?¹¹

⁹ Here, financial independence is measured as receiving less than 50% of the core funding for basic educational services from government agencies.

¹⁰ For methodological details, cf. Wößmann (2003a, 2003b) and Fuchs and Wößmann (2006).

¹¹ The results are only briefly summarized here. For considerably more detail, cf. Wößmann (2002, 2003a) for the results using TIMSS data, Wößmann (2003b, 2003c) for TIMSS-

3.1 Competition from Privately Managed Schools

The first institutional feature analyzed is the availability of privately managed schools, which provide competition for public schools and choice for parents. Economic models of industrial organization suggest that competition and choice create incentives that further performance. Theoretical applications to the market for education are numerous, often with differing focuses and conflicting predictions in terms of distributional consequences.¹² However, the basic thrust of these models in terms of the efficiency of the education system is that choice and competition in education can create incentives for cost containment and performance-conducive qualitative innovation, as customers (parents) get involved in choosing those suppliers that promise best performance.

The bottom line of the evidence from international achievement tests on competition from private schools is that students perform better in countries where more schools are privately managed. For example, students scored 10 test-score points better in TIMSS math, and 9 in science, if the share of enrollment in privately managed schools of a country was 1 international standard deviation (or 14 percentage points) higher (cf. Wößmann, 2003a).¹³ Considering that one grade-level equivalent (the average performance difference between 7th and 8th grade) on average was roughly equal to 40 test-score points on the TIMSS test, this is a very large effect indeed. Put differently, students in countries that had a private school sector that was 28 percentage points larger (as measured by the enrollment share) on average performed better by the equivalent of half a year's learning.

In addition to private enrollment, students in countries with a higher share of public educational spending going to private institutions performed better. If the share of public funds going to independent private schools rose by 1 percentage point (or 1 international standard deviation), there was a 10 test-score point increase in math achievement. In sum, student performance seems to be higher in education systems where taxpayers' money is allocated by private schools rather than by the public schooling system.

The evidence discussed so far, using TIMSS data, is based on country-wide measures of the extent of private schooling. This does not allow for a direct assessment of the relative performance of public and private schools, because TIMSS does not provide school-level data on whether individual tested schools are public or private. However, measuring the system-level effect of private school

Repeat, Fuchs and Wößmann (2006) and Wößmann (2005d) for PISA, Wößmann (2005a) for all three and Fuchs and Wößmann (2004b) for PIRLS.

¹² Cf., e.g., Chubb and Moe (1990), Shleifer (1998), Epple and Romano (1998), Nechyba (2000) and Gradstein et al. (2004).

¹³ These results refer to the OECD countries participating in TIMSS, for whom consistent data on the share of private schools are available.

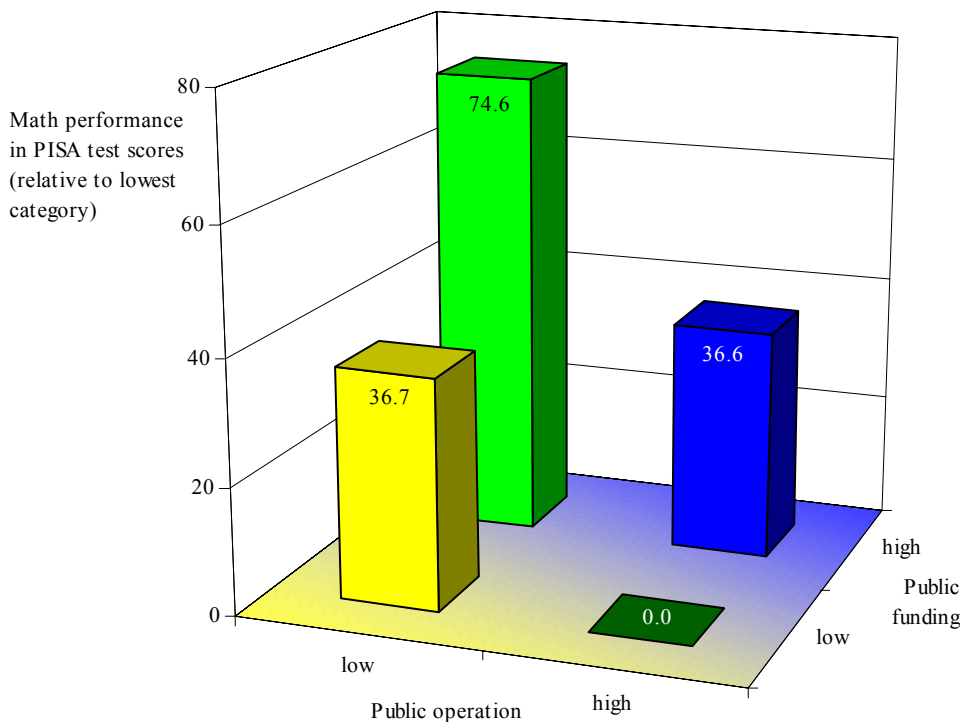
management may be the appropriate way to estimate the general systemic effect of the competitive environment prevailing in the different education systems, because increased competition from private schools may also positively impact on the effectiveness of resource use in nearby public schools.

By contrast, PISA for the first time provides specific school-level data on public versus private management and financing. In particular, in PISA there is information for each tested school both on whether the school is privately or publicly managed and on how large its share of public funding is. Public school management is defined as schools managed directly or indirectly by a public education authority, government agency or governing board appointed by government or elected by public franchise, whereas private school management is defined as schools managed directly or indirectly by a non-government organization, for example a church, trade union, businesses or other private institutions. The share of public funding is defined as the percentage of total school funding coming from government sources (at different levels), as opposed to such private contributions as fees and donations.

Looking across all countries, the result is that students perform better if their specific school is privately managed. The size of the performance difference between privately and publicly managed schools is between 16 and 20 PISA test-score points in the three different subjects (Fuchs and Wößmann, 2006). When interpreting these results based on micro-level variations within countries, one should be cautious, though, because there may be self-selection of students with different capabilities into private versus public schools. While many features of self-selection will be held constant by the extensive family-background controls that the analyzes contain, there cannot be final confidence about whether some self-selection bias remains due to unobserved heterogeneity of students.

Wößmann (2005d) provides a more in-depth analysis of the effects of private vs. public management and financing of schools in PISA, mostly measuring these features at the country level. This approach allows to capture systemic effects where both private and public schools may perform at a higher level because of the existence of private competition. By contrast, if public schools behave differently because there are private schools nearby, then there may be effects of private involvement even though the performance between individual private and public schools may not differ at the level of schools. The results show again that countries with a larger share of privately managed schools perform better. At the same time, across countries, larger shares of public funding (as opposed to management) are associated with better student outcomes. This pattern is depicted in chart 4, which shows that countries which combine relatively high shares of private operation with relatively high shares of public funding do best among all possible operation-funding combinations, while countries which combine public operation with private funding do worst.

Chart 4: Effects of Private versus Public Management and Financing of Schools



Source: Wößmann (2005d).

Furthermore, at the school level the advantage of privately operated schools over publicly operated schools is particularly strong in countries with large shares of public funding. This suggests that public funding may increase the set of choices of poor families. Without public funding, poor families may be constrained in their choices because they do not have the financial means to opt for private schooling. In these cases, public funding may help families to exert their choices in terms of privately managed schools, so that the positive effect of public funding may be another aspect of the skill-enhancing capacity of school choice and competition. Keeping the caveat in mind that studies based on observational data have limits in terms of causal interpretations, the international evidence suggests that school systems based on public-private partnerships where the state finances schools but contracts their operation out to the private sector seem to be the most effective school systems in terms of fostering students' educational performance.

This evidence on positive performance effects of school choice from the international tests is consistent with other evidence from within countries.¹⁴ The evidence provided by Neal (1997) suggests that the choice of private Catholic schools leads to higher performance of inner-city students in the United States. Hoxby (2003b) summarizes ample evidence from recent policy experiments in the United States showing that school choice and school competition, among others in the form of vouchers and charter schools (relatively autonomous public schools that give parents additional choice), improve the performance not only of these schools, but also of the public schools that face their competition. Howell et al. (2002) provide evidence from several randomized field trials in the United States showing that school vouchers substantially increased the academic performance of African Americans who were enabled to switch to a private schools. Within the system of public schools, increased competition among U.S. public schools has also been shown to improve student performance (Hoxby, 2000).

Outside the U.S.A., Bradley and Taylor (2002) and Levačić (2004) find similar positive effects of school competition on the performance of English schools. Sandström and Bergström (2005) and Björklund et al. (2004) provide evidence on significant positive effects of competition from privately operated schools on the performance of public schools in Sweden. Filer and München (2003) show that the introduction of a voucher-type system in the Czech Republic led to the creation of private schools in areas where public schools are doing badly and that the public schools facing private competition improved their performance in obtaining university admission for their graduates. The benefits of a program that provided vouchers for the attendance of private schools in Colombia have been found to clearly exceed its cost, which was similar to providing a place in public schools (Angrist et al., 2002).

All this shows that competition from private schools can have positive effects on students' academic achievement. Obviously, there are also important caveats to keep in mind with implementing competition in the education field. Critics particularly fear sorting and adverse effects on disadvantaged students (e.g., Burgess et al., 2006; Cullen et al., 2003; Ladd 2002), although the evidence sometimes even points in the opposite direction of equalizing effects (e.g., Hoxby, 2003b; Nechyba, 2000). Others argue that a universal voucher system may bear considerable administrative costs (Levin, 1998). While much more research is needed before we understand fully the working of competition in education and the circumstances which determine its effects, the available evidence strongly suggests that the use of competition from private educational providers, combined with public funding of schools, can increase the efficiency with which students receive necessary competencies.

¹⁴ Cf. Hoxby (2003a) for a collection of recent research on the economics of school choice.

3.2 School Autonomy

A second set of institutional features analyzed is the extent of autonomy that schools have, depicting the extent to which schools and teachers can make their own choices. Economic models of the centralization or decentralization of school operation suggest that larger autonomy can lead to increased efficiency of public schools (cf., e.g., Bishop and Wößmann, 2004).

The general pattern of results on school autonomy from the international tests is that students perform better in schools that have autonomy in process and personnel decisions (Wößmann, 2003a; Fuchs and Wößmann, 2006). These decisions include such areas as deciding on the purchase of supplies and on budget allocations within schools, hiring and rewarding teachers (within a given budget) and choosing textbooks, instructional methods etc. That is, there are positive performance effects of choice for schools in these specific decision-making areas. This general result is found both in the secondary- and in the primary-school international tests (Fuchs and Wößmann, 2004b).

Similarly, students perform better if their teachers have both incentives and powers to select appropriate teaching methods (Wößmann, 2003a; Fuchs and Wößmann, 2006). In this sense, there are also positive performance effects of choice for teachers – as long as they are held accountable for what they do (see section 3.4 below).

3.3 Accountability through External Exams

Principal-agent models of educational production predict that setting clear performance standards and providing performance information can tilt incentives in favor of superior student performance (cf., e.g., Costrell, 1994; Betts, 1998). In particular, by signaling student performance to potential employers on the labor market, external school-leaving exams increase students' rewards for learning as well as parents' scope for monitoring the education process, which should ultimately improve student performance (cf., e.g., Bishop and Wößmann, 2004; Bishop, 2006). The accountability introduced by external exams can help to face the challenge for the institutional set-up of school systems to create a set of incentives that encourages school personnel to behave in ways that do not necessarily further their own interests, but rather the interest of best student learning. For instance, without the right incentives, teachers may avoid using the most promising teaching techniques, preferring to use the techniques they find most convenient. If a country assesses the performance of students with some sort of external exam and uses this information to monitor teachers, teachers may put aside their other interests and focus mainly on raising student achievement. In sum,

testing performance can make students and educational providers accountable for what they learn and teach.¹⁵

The evidence from the international student achievement tests shows exactly that. Students perform substantially better in countries that have external exit-exam systems than in countries without external exit-exam systems. This is true in TIMSS, in TIMSS-Repeat and in PISA (cf. Wößmann, 2003a, 2003b, 2005a; Fuchs and Wößmann, 2006), as well as in other previous international achievement tests (cf. also Bishop, 1997, 2006). By and large, the evidence suggests that the effect may well be larger than a whole grade-level equivalent. That is, student performance is immensely higher where schools and students are held accountable by external exams.

Similarly, students perform better where parents take interest in teaching matters, suggesting positive effects both of parental choices and of parents holding schools and children accountable (Wößmann, 2003a). Also, students perform better where teachers place a lot of emphasis on monitoring student progress by regular tests and exams (Fuchs and Wößmann, 2006). This is additional evidence that accountability for students increases their educational performance. Furthermore, this is the case in primary school (PIRLS) just as well as in secondary school (Fuchs and Wößmann, 2004b).

In the two national education systems where the existence of external exams varies within countries because some regions feature them and others not, Canada and Germany, it has similarly been shown that students perform better in regions with external exams (cf. Bishop 1997; Jürges et al., 2005). In a related literature, Figlio and Lucas (2004) report U.S. evidence on positive effects of grading standards on student achievement. Another means to increase accountability are explicit school-focused accountability systems, which have been shown to increase students' learning achievement in the United States (Carnoy and Loeb, 2003; Hanushek and Raymond, 2004; Jacob, 2005). One institutional set-up that combines accountability with parental choice are systems which give students in schools that repeatedly do badly on the accountability test a voucher to attend private schools. In Florida, the threat of becoming subject to private-school choice if failing on the test has been shown to increase school performance particularly for disadvantaged students (West and Peterson, 2006).

It should be borne in mind, though, that designing proper accountability systems that hold actors accountable for only those outcomes for which they are really responsible is not an easy task. External exit examinations can introduce incentives for students if they produce signals of accomplishment that have real consequences for students. Bishop (2006) suggests that a well-designed system of external exit examinations should be curriculum-based, define achievement relative to an

¹⁵ Two recent collections of work on accountability are Evers and Walberg (2002) and Peterson and West (2003).

external standard, measure the full range and signal multiple levels of achievement, and cover the vast majority of students.

By contrast, accountability systems that aim to create proper incentives for schools require a value-added approach which tests the learning gains (rather than levels) of each individual student (cf. Kane and Staiger, 2002; Ladd and Walsh, 2002). School-focused accountability systems can also lead to strategic responses on part of teachers and schools, for example by increasing placements of low-performing students in special-education programs which are outside the accountability system or by pre-emptively retaining students (Jacob, 2005). High-stakes testing may also introduce incentives for cheating (Jacob and Levitt, 2003). Thus, in implementing accountability systems, it is crucial to provide means that keep strategic responses and fraud to a minimum. By contrast, worries about the direct costs of implementing accountability systems should not be overstated, as the costs of the accountability programs implemented in several U.S. states that include comprehensive external testing have been shown to be minuscule (Hoxby, 2002).

3.4 External Exams as the “Currency” of Decentralized School Systems

So far, school autonomy and external exams were considered as unrelated institutional features. However, there are reasons to expect that external exams and school autonomy are complementary, in the sense that the one is particularly effective if the other is also in place (cf. Wößmann, 2005a for details). Put differently, external exams are a pre-requisite for decentralized, choice-based systems of autonomous schools to function properly. In this sense, external exams are the “currency” of decentralized school systems (Wößmann, 2003c).

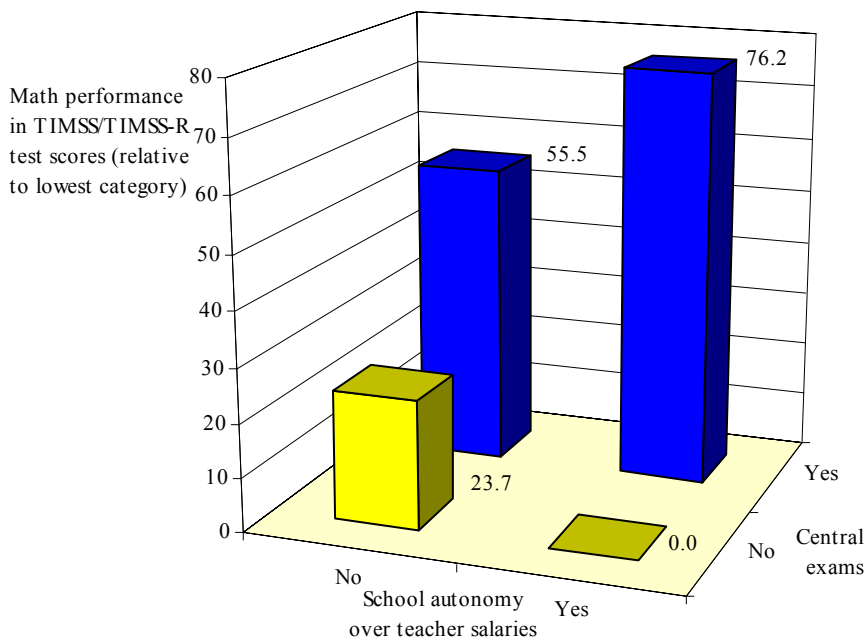
In the economic system, money is an institutional feature that allows one to value and compare different objects. This kind of price system creates knowledge that no single person can gather. External exams can provide such “price information” to the education system. The important feature is that the exams are instituted as standardized tests by independent institutions and in a manner external to the individual school, so that they provide independent and comparable information on how the school performs. Parents can use this information created by external exams to make proper choices. This is the core of the idea of accountability: It creates competition where beforehand no comparable yardstick was available to make informed choices. Once this “price system” is in place, a system of decentralized, autonomous schools can be expected to work much better than any centralized system could, both because the autonomous schools can use their superior local knowledge about how to best teach their students and because competition provides them incentives to focus their efforts on student learning.

This assertion can be corroborated by evidence from the cross-country pattern of student performance. The results show that external exit exams improve educational performance, and at the same time that school autonomy is more beneficial in systems with external exams (Wößmann, 2003c, 2005a; Fuchs and Wößmann, 2006). In several decision-making areas, external exams even turn an initially negative autonomy effect around into a positive effect.

One such case is depicted in chart 5, which plots students' math performance in TIMSS and TIMSS-Repeat under the four conditions resulting from the presence and absence of central exams and school autonomy over teacher salaries: The performance of students in schools without salary autonomy in systems without central exams; with autonomy but without central exams; without autonomy but with central exams; and with both autonomy and central exams. Performance is depicted relative to the condition with the lowest performance, which is the condition of salary autonomy without central exams.

As chart 5 shows, school autonomy regarding teacher salaries has a *negative* effect on student performance in systems without central exams. In systems with central exams, student performance is generally higher than in systems without central exams, both in the case with and without school autonomy. In addition, however, it is striking that the effect of school autonomy is turned completely around in systems with central exams: Salary autonomy of schools has *positive* effects on student performance in central-exam systems.

This is strong evidence of complementarity between accountability and decentralized choice. Without the accountability introduced by central exams, schools behave opportunistically because their local opportunistic behavior cannot be externally observed and thus cannot be sanctioned. Hence school decision-makers do not feel obliged to set teacher salaries so as to contribute to enhancing student performance, but can use their decision-making autonomy to promote other interests. In contrast, central exams provide information about whether the schools perform well or not, so that parents and supervisory authorities can draw possible consequences from school behavior that weakens performance. This creates incentives for decision-makers in schools not to exploit their autonomy in setting teacher salaries in an opportunistic way, but to use it in order to effectively promote student performance. The benefits of superior local knowledge then come into effect, as school decision-makers ought to know better than any central authority which teachers deserve to be rewarded for good work.

Chart 5: External Exams, Salary Autonomy and Learning

Source: Wößmann (2005a).

That is, the accountability introduced by the “price information” of external exams creates competition, which brings the beneficial effects of local school choices to the fore. The very same effects of school autonomy over teacher salaries with and without central exams are found not only in TIMSS and TIMSS-Repeat, but also in PISA (Fuchs and Wößmann, 2006). Likewise, similar cases where external exams turn a negative autonomy effect around into a positive effect have been found for such decision-making areas as school autonomy in determining course content and teacher influence on resource funding. More generally, in several additional decision-making areas the general pattern of the evidence suggests that school autonomy is better for student performance when external exit exams are in place (cf. Wößmann, 2005a for details).

In sum, external exams can be regarded as the “currency” of school systems: They are a measure of value which prevents decentralized opportunism. As such, they are a precondition for decentralized education systems to achieve high student performance. Efficient education policies would thus combine external exams with school autonomy, setting and testing standards externally but leaving it up to schools how to pursue them.

4. Summary and Conclusion

Without doubt, fostering the human capital of the population will have to be part of any successful growth strategy in today's knowledge-based economy. But how can we improve the quality of schools that produce this human capital? The conclusion that can be drawn from the evidence based on four extensive international student achievement tests is clear: Institutions matter! In particular, institutional features that ensure choice, autonomy and accountability in the school system are key to high student performance. The different institutional effects add up to a huge aggregate effect. For example, their effects in TIMSS add up to more than 200 test-score points, compared to an international standard deviation of 100 test-score points and to a grade-level equivalent of 40 test-score points (Wößmann, 2003a). Similarly, about a quarter of the total international variation in educational performance in PISA can be accounted for by international variation in the institutional features (Fuchs and Wößmann, 2006). That is, the institutional effects are very large indeed.

The lessons that school policy can learn from the cross-country evidence include that students perform better:

- in countries with more *competition* from privately managed schools;
- in countries where *public funding* ensures that all families can make choices;
- in *schools* that have *autonomy in process and personnel decisions*;
- if their *teachers* have both incentives and power to select appropriate *teaching methods*;
- where *parents* take *interest* in teaching matters;
- where student progress is monitored by regular *testing*;
- where *schools* are held *accountable by external exams*; and
- where external exams and school autonomy are combined.

The evidence based on international comparisons across numerous countries allows all countries to learn from each other in terms of what works best in the education system. No single country in the world has the single “first best” education system that does everything right. The cross-country perspective taken in this paper enables the exploitation of institutional variations between all the participating countries. Thereby, it allows both to analyze the underlying reasons for differing performance and to learn from each other in terms of revealed best educational practice.

It is clear that this international evidence can only provide the “big picture” of results, revealing broad patterns but not specifics of implementation details. Surely, implementation is crucial with any of the institutional features discussed, and more detailed research is needed to learn how to implement competition, autonomy and accountability in different circumstances. But by depicting the “average” effect of these institutions as implemented in the real-world education systems across the

countries, the cross-country results can reveal some of the main driving forces of success in the education system.

Also, looking at competition, autonomy and accountability is not an exhaustive treatment of the relevant institutional features of education systems. For example, monetary incentives for teachers based on their students' performance have been shown to improve student learning in Israel immensely (Lavy, 2002, 2004). Similarly, Atkinson et al. (2004) find that the introduction of performance-related pay had a substantial positive impact on student achievement in England.¹⁶ Teacher incentives are particularly crucial because arguably, apart from the students themselves, teachers constitute the most important "input" in the education process, in terms of both cost and content (cf. Rivkin et al., 2005). Another institutional feature with possibly important implications for educational performance is the extent of tracking of students into different types of school, which has been shown to be associated with increased inequality of student achievement across countries (Hanushek and Wößmann, 2006). Likewise, the extent of the pre-school education system can have large impact on students' later learning achievement (Schütz et al., 2005). Thus, interventions at early ages may be particularly relevant, given the importance of early childhood investments for later human capital investments over the life cycle (Carneiro and Heckman, 2003; Cunha et al., 2006).

When asking how education policies can create the competencies and learning achievements required for citizens and societies to prosper in the future, the binding constraint seems to be institutional reforms, not resource expansions within the current institutional systems. For educational investments to translate into student learning, all the people involved in the education process have to face the right incentives that make them act in ways that advance student performance. The international evidence summarized in this paper suggests that institutional structures that create performance-conducive incentives by introducing competition, autonomy and accountability stand a good chance of improving the quality of schools which is crucial for long-run economic growth.

References

- Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, Michael Kremer (2002). Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment. *American Economic Review* 92 (5): 1535–1558.
- Atkinson, Adele, Simon Burgess, Bronwyn Croxson, Paul Gregg, Carol Propper, Helen Slater, Deborah Wilson (2004). Evaluating the Impact of Performance-

¹⁶ Atkinson et al. (2004) provide a survey of additional studies on performance-related teacher pay, the more rigorous of which also tend to find a positive relationship between financial teacher incentives and student outcomes.

- related Pay for Teachers in England. CMPO Working Paper 04/113. Bristol: Centre for Market and Public Organisation.
- Barro, Robert J. (2001). Human Capital and Growth. *American Economic Review, Papers and Proceedings* 91 (2): 12–17.
- Betts, Julian R. (1998). The Impact of Educational Standards on the Level and Distribution of Earnings. *American Economic Review* 88 (1): 266–275.
- Bishop, John H. (1997). The Effect of National Standards and Curriculum-Based Examinations on Achievement. *American Economic Review* 87 (2): 260–264.
- Bishop, John H. (2006). Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn. Forthcoming in: Eric A. Hanushek, Finis Welch (eds.), *Handbook of the Economics of Education*, Amsterdam: North-Holland.
- Bishop, John H., Ludger Wößmann (2004). Institutional Effects in a Simple Model of Educational Production. *Education Economics* 12 (1): 17–38.
- Björklund, Anders, Per-Anders Edin, Peter Freriksson, Alan Krueger (2004). Education, Equality and Efficiency: An Analysis of Swedish School Reforms during the 1990s. IFAU Report 2004:1. Uppsala: Institute for Labour Market Policy Evaluation.
- Bradley, Steve, Jim Taylor (2002). The Effect of the Quasi-Market on the Efficiency-Equity Trade-off in the Secondary School Sector. *Bulletin of Economic Research* 54 (3): 295–314.
- Burgess, Simon, Brendon McConnell, Carol Propper, Deborah Wilson (2006). The Impact of School Choice on Sorting by Ability and Socio-economic Factors in English Secondary Education. Forthcoming in: Paul E. Peterson, Ludger Wößmann (eds.), *Schools and the Equal Opportunity Problem*. Cambridge, MA: MIT Press.
- Carneiro, Pedro, James J. Heckman (2003). Human Capital Policy. In: James J. Heckman, Alan B. Krueger, *Inequality in America: What Role for Human Capital Policies?*. Cambridge, MA: MIT Press (also available as NBER Working Paper 9495, Cambridge, MA: National Bureau of Economic Research).
- Carnoy Martin, Susanna Loeb (2003). Does External Accountability Improve Student Outcomes? A Cross-State Analysis. *Education Evaluation and Policy Analysis* 24 (4): 305–331.
- Chubb, John E., Terry M. Moe (1990). *Politics, Markets, and America's Schools*. Washington, D.C.: Brookings Institution Press.
- Costrell, Robert M. (1994). A Simple Model of Educational Standards. *American Economic Review* 84 (4): 956–971.
- Cullen, Julie B., Brian A. Jacob, Steven D. Levitt (2005). The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools. *Journal of Public Economics* 89 (5–6): 729–760.
- Cunha, Flavio, James J. Heckman, Lance Lochner, Dimitriy V. Masterov (2006). Interpreting the Evidence on Life Cycle Skill Formation. Forthcoming in: Eric

- A. Hanushek, Finis Welch (eds.), *Handbook of the Economics of Education*. Amsterdam: North-Holland. (available as NBER Working Paper 11331, Cambridge, MA: National Bureau of Economic Research, 2005)
- De la Fuente, Angel (2004). Human Capital and Growth: Some Results for the OECD. In: Oesterreichische Nationalbank (ed.), *Current Issues of Economic Growth*, Proceedings of OeNB Workshops No. 2, Vienna: Oesterreichische Nationalbank.
- Epple, Dennis, Richard E. Romano (1998). Competition between Private and Public Schools, Vouchers, and Peer-Group Effects. *American Economic Review* 88 (1): 33–62.
- Evers, Williamson M., Herbert J. Walberg (eds.) (2002). *School Accountability*. Stanford, CA: Hoover Institution Press.
- Figlio, David N., Maurice E. Lucas (2004). Do High Grading Standards Affect Student Performance? *Journal of Public Economics* 88 (9): 1815–1834.
- Filer, Randall K., Daniel Münich (2003). Responses of Private and Public Schools to Voucher Funding. Paper presented at the Annual Meeting of the American Economic Association in Washington, D.C.
- Fuchs, Thomas, Ludger Wößmann (2004a). Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and Use of Computers at Home and at School. *Brussels Economic Review* 47 (3/4): 359–385.
- Fuchs, Thomas, Ludger Wößmann (2004b). The Determinants of Differences in Primary-School Learning across Countries. Mimeo, Ifo Institute for Economic Research at the University of Munich.
- Fuchs, Thomas, Ludger Wößmann (2006). What Accounts for International Differences in Student Performance? A Re-examination Using PISA Data. *Empirical Economics*: forthcoming. (available as CESifo Working Paper 1235, Munich: CESifo, 2004)
- Gradstein, Mark, Moshe Justman, Volker Meier (2004). *The Political Economy of Education: Implications for Growth and Inequality*. Cambridge, MA: MIT Press.
- Gundlach, Erich, Desmond Rudman, Ludger Wößmann (2002). Second Thoughts on Development Accounting. *Applied Economics* 34 (11): 1359–1369.
- Gundlach, Erich, Ludger Wößmann, Jens Gmelin (2001). The Decline of Schooling Productivity in OECD Countries. *Economic Journal*. 111 (471): C135–C147.
- Hanushek, Eric A. (2003). The Failure of Input-Based Schooling Policies. *Economic Journal* 113 (485): F64–F98.
- Hanushek, Eric A. (2005). The Economics of School Quality. *German Economic Review* 6 (3): 269–286.
- Hanushek, Eric A., with others (1994). *Making Schools Work: Improving Performance and Controlling Costs*. Washington, D.C.: Brookings Institution Press.

- Hanushek, Eric A., Dennis D. Kimko (2000). Schooling, Labor-Force Quality, and the Growth of Nations. *American Economic Review* 90 (5): 1184–1208.
- Hanushek, Eric A., Margaret E. Raymond (2004). The Effect of School Accountability Systems on the Level and Distribution of Student Achievement. *Journal of the European Economic Association* 2 (2–3): 406–415.
- Hanushek, Eric A., Ludger Wößmann (2006). Does Early Tracking Affect Educational Inequality and Performance? Difference-in-Difference Evidence across Countries. *Economic Journal* 116 (510): C63–C76.
- Howell, William G., Patrick J. Wolf, David E. Campbell, Paul E. Peterson (2002). School Vouchers and Academic Performance: Results from Three Randomized Field Trials. *Journal of Policy Analysis and Management* 21 (2): 191–217.
- Hoxby, Caroline M. (2000). Does Competition among Public Schools Benefit Students and Taxpayers? *American Economic Review* 90 (5): 1209–1238.
- Hoxby, Caroline M. (2002). The Cost of Accountability. In: Williamson M. Evers, Herbert J. Walberg (eds.), *School Accountability*. Stanford, CA: Hoover Institution Press.
- Hoxby, Caroline M. (ed.) (2003a). *The Economics of School Choice*. A National Bureau of Economic Research Conference Report. Chicago: University of Chicago Press.
- Hoxby, Caroline M. (2003b). School Choice and School Competition: Evidence from the United States. *Swedish Economic Policy Review* 10 (3): 9–65.
- Jacob, Brian A. (2005). Accountability, Incentives and Behavior: The Impact of High-stakes Testing in the Chicago Public Schools. *Journal of Public Economics* 89 (5–6): 761–796.
- Jacob, Brian A., Steven D. Levitt (2003). Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating. *Quarterly Journal of Economics* 118 (3): 843–877.
- Jürges, Hendrik, Kerstin Schneider, Felix Büchel (2005). The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany. *Journal of the European Economic Association* 3 (5): 1134–1155.
- Kane, Thomas J., Douglas O. Staiger (2002). The Promise and Pitfalls of Using Imprecise School Accountability Measures. *Journal of Economic Perspectives* 16 (4): 91–114.
- Krueger, Alan B. (2003). Economic Considerations and Class Size. *Economic Journal* 113 (485): F34–F63.
- Ladd, Helen F. (2002). School Vouchers: A Critical View. *Journal of Economic Perspectives* 16 (4): 3–24.
- Ladd, Helen F., Randall P. Walsh (2002). Implementing Value-Added Measures of School Effectiveness: Getting the Incentives Right. *Economics of Education Review* 21 (1): 1–17.

- Lavy, Victor (2002). Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement. *Journal of Political Economy* 110 (6): 1286–1317.
- Lavy, Victor (2004). Performance Pay and Teachers' Effort, Productivity and Grading Ethics. NBER Working Paper 10622. Cambridge, Mass.: National Bureau of Economic Research.
- Levačić, Rosalind (2004). Competition and the Performance of English Secondary Schools: Further Evidence. *Education Economics* 12 (2): 177–193.
- Levin, Henry M. (1998). Educational Vouchers: Effectiveness, Choice, and Costs. *Journal of Policy Analysis and Management* 17 (3): 373–392.
- Neal, Derek (1997). The Effects of Catholic Secondary Schooling on Secondary Achievement. *Journal of Labor Economics* 15 (1): 98–123.
- Nechyba, Thomas J. (2000). Mobility, Targeting, and Private-School Vouchers. *American Economic Review* 90 (1): 130–146.
- Peterson, Paul E., Martin R. West (eds.) (2003). *No Child Left Behind? The Politics and Practice of School Accountability*. Washington, D.C.: Brookings Institution Press.
- Pritchett, Lant (2002). "When Will They Ever Learn"? Why All Governments Produce Schooling. Mimeo, Harvard University, Kennedy School of Government.
- Rivkin, Steven G., Eric A. Hanushek, John F. Kain (2005). Teachers, Schools, and Academic Achievement. *Econometrica* 73 (2): 417–458.
- Sandström, F. Mikael, Fredrik Bergström (2005). School Vouchers in Practice: Competition Will Not Hurt You. *Journal of Public Economics* 89 (2–3): 351–380.
- Schütz, Gabriela, Heinrich W. Ursprung, Ludger Wößmann (2005). Education Policy and Equality of Opportunity. CESifo Working Paper 1518. Munich: CESifo.
- Shleifer, Andrei (1998). State versus Private Ownership. *Journal of Economic Perspectives* 12 (4): 133–150.
- West, Martin R., Paul E. Peterson (2006). The Efficacy of Choice Threats within School Accountability Systems: Results from Legislatively-Induced Experiments. *Economic Journal* 116 (510): C46–C62.
- Wößmann, Ludger (2002). *Schooling and the Quality of Human Capital*. Berlin: Springer.
- Wößmann, Ludger (2003a). Schooling Resources, Educational Institutions and Student Performance: the International Evidence. *Oxford Bulletin of Economics and Statistics* 65 (2): 117–170.
- Wößmann, Ludger (2003b). Central Exit Exams and Student Achievement: International Evidence. In: Paul E. Peterson, Martin R. West (eds.), *No Child Left Behind? The Politics and Practice of School Accountability*. Washington, D.C.: Brookings Institution Press.

- Wößmann, Ludger (2003c). Central Exams as the “Currency” of School Systems: International Evidence on the Complementarity of School Autonomy and Central Exams. *DICE Report – Journal for Institutional Comparisons* 1 (4): 46–56.
- Wößmann, Ludger (2003d). Specifying Human Capital. *Journal of Economic Surveys* 17 (3): 239–270.
- Wößmann, Ludger (2004). Institutional Comparisons in Educational Production. *CESifo DICE Report – Journal for Institutional Comparisons* 2 (4): 3–6.
- Wößmann, Ludger (2005a). The Effect Heterogeneity of Central Exams: Evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics* 13 (2): 143–169.
- Wößmann, Ludger (2005b). Evidence on the Effects of Choice and Accountability from International Student Achievement Tests. In: David Salisbury, James Tooley (eds.), *What Americans Can Learn from School Choice in Other Countries*, pp. 133–148. Washington, D.C.: Cato Institute.
- Wößmann, Ludger (2005c). Educational Production in Europe. *Economic Policy* 20 (43): 445–504.
- Wößmann, Ludger (2005d). Public-Private Partnerships in Schooling: Cross-Country Evidence on their Effectiveness in Providing Cognitive Skills. Program on Education Policy and Governance, Research Paper PEPG 05–09. Cambridge, MA: Harvard University.
- Wößmann, Ludger, Martin R. West (2006). Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS. *European Economic Review* 50 (3): 695–736.