Open access • Journal Article • DOI:10.1111/1468-0084.00045

Schooling Resources, Educational Institutions and Student Performance: the International Evidence — Source link

Ludger Wößmann

Institutions: Ifo Institute for Economic Research

Published on: 01 May 2003 - Oxford Bulletin of Economics and Statistics (Blackwell Publishing Ltd)

Topics: Curriculum

Wößmann, Ludger

Working Paper

Schooling Resources, Educational Institutions, and Student Performance: The International Evidence

Kiel Working Paper, No. 983

Provided in Cooperation with:
Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

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

Kiel Working Paper No. 983

Schooling Resources, Educational Institutions, and Student Performance: The International Evidence

by

Ludger Wößmann

Revised Version
December 2000

Schooling Resources, Educational Institutions, and Student Performance: The International Evidence

Abstract:

The paper suggests that international differences in educational institutions explain the large international differences in student performance in cognitive achievement tests. A microeconometric student-level estimation based on data for more than 260,000 students from 39 countries reveals that positive effects on student performance stem from centralized examinations and control mechanisms, school autonomy in personnel and process decisions, competition from private educational institutions, scrutiny of achievement, and teacher influence on teaching methods. A large influence of teacher unions on curriculum scope has negative effects on student performance. The findings imply that international differences in student performance are not caused by differences in schooling resources but are mainly due to differences in educational institutions.

Keywords: education production function, institutions, incentives

JEL Classification: I2

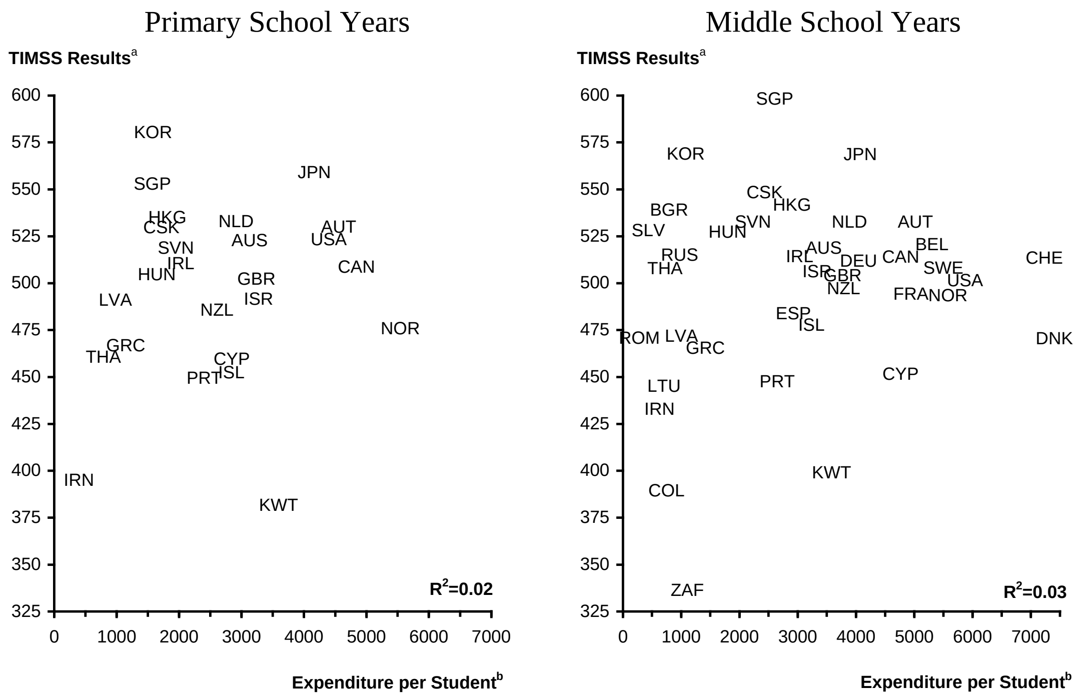
Ludger Wößmann
Kiel Institute of World Economics
24100 Kiel, Germany
Telephone: (+49) 431 8814-497
Fax: (+49) 431 8814-500
E-mail: woessmann@ifw.uni-kiel.de

1 Introduction and Summary

The formation of human capital is essential for the economic success both of individuals and of society at large in a modern economy. The human capital stock comprises cognitive and non-cognitive skills and is mainly produced in families, schools, universities, and firms. This study focuses on students' cognitive skills in mathematics and science, which are mainly formed in schools. Since "[e]arly learning begets later learning" (Heckman 1999, p. 2), basic knowledge formed early in school has a substantial impact on potential future prosperity of individuals and nations.

The empirical evidence on the determinants of educational performance overwhelmingly shows that at given levels of expenditures, an increase in the amount of resources used does not generally lead to an increase in educational performance. The lack of a strong and systematic relationship between resources and performance has been shown within the United States (Hanushek 1986, 1996), within developing countries (Hanushek 1995), across countries (Hanushek and Kim 1999), and across time within most OECD countries (Gundlach et al. 2001) and within some East Asian countries (Gundlach and Wößmann 1999). Figure 1 presents equivalent evidence from the latest and most coherent cross-country achievement study, the Third International Mathematics and Science Study (TIMSS). Again, cross-country differences in expenditure per student do not help in understanding cross-country differences in educational performance. The correlation coefficient between expenditure per student and average TIMSS test scores is 0.13 in the primary school years and 0.16 in the middle school years. By implication, the level of schooling productivity - the ratio of educational performance to resources used - seems to differ widely across different schooling systems.

Figure 1: Expenditure per student and educational performance: the cross-country evidence



a Average of mathematics and science results in 3rd/4th grade and in 7th/8th grade, respectively.
b At primary and secondary level, respectively, in international dollars, 1994.

Sources: IEA (1998), UNESCO (var. iss.).

An extensive debate has unfolded whether the absence of any statistically significant positive input-output relation can be taken at face value. Critics point to problems in the conduct of meta-analyses (Hedges et al. 1994), to the use of alternative output measures like labor market performance (Card and Krueger 1992, Heckman et al. 1996, Case and Yogo 1999), to experimental estimates (Krueger 1999), and to quasi-controlled empirical experiments showing positive resource effects (Angrist and Lavy 1999, Case and Deaton 1999).[1] Notwithstanding this debate and the fact that there certainly are circumstances where resources do matter, the large

1 For discussions of the effectiveness of resource use in education, see also JEP 1996, RES 1996, and FRBNY 1998.

international differences in student performance levels in mathematics and science are a fact, and it is obvious and generally accepted that differences in the amount of inputs used do not suffice as an explanation of their occurrence. The TIMSS micro data base begs a wealth of information to understand these differences, because it is based on performance tests of individual students in about 40 countries which can be combined with student-specific information about the characteristics and influences of students, parents, teachers, schools, and the administration.

It has been argued that public schooling systems do not set suitable incentives for improving students' performance or containing costs (Hanushek et al. 1994). As the Economist (January 16th, 1999, p. 21) put it, "[i]n most countries the business of running schools is as firmly in the grip of the state as was the economy of Brezhnev's Russia." While public provision of schooling may generally be associated with inefficiencies, the public schooling systems still differ substantially across countries in their institutional structure of educational decision-making processes. They give different amounts of decision-making power to the different agents involved in educational production, which creates different incentives for their behavior. These differences in institutions and incentives will affect the agents' decisions on resource allocation and thereby the effectiveness of resource use in the education sector, which should impact on the educational performance of the students. This paper examines whether and, if so, how differences in these institutional incentive mechanisms can add to an explanation of the large international differences in students' cognitive skills.

In the education process, a network of principal-agent relationships exists which entail conflicts between the multifarious interests of different groups and serious problems of monitoring due to informational advantages of self-interested agents (Section 2). This can create adverse incentives and leeway for the agents to act opportunistically, leading to an inefficient use of given resources and to misallocations of resources across different uses. Institutional features such as external examinations and a competitive environment set by a large private schooling sector should focus agents' interests on students' learning and should be best suited to face monitoring problems, thereby being conducive to student performance. In a similar way, the distribution of decision-making powers and responsibilities between schools and the educational administration, as well as between different administrative levels, should impact on the prevailing incentive structure and on educational outcomes. Institutions determining the influence, freedom of action, and accountability of teachers, students, and parents should also matter for the agents' incentives and for student performance.

However, many of these relationships are very complex, and the interests of the different agents are far from being unidimensional. Therefore, an empirical investigation seems the best step forward to determine the direction and strength of the impact of different institutions and to enhance our understanding of the determinants of student performance. The link between institutions and student performance can be tested by estimating an education production function which includes data on institutions as explanatory variables. Since there is no significant variation in many institutional features within a single country on which such an analysis could be based, country-specific evidence can hardly be used to perform this test. Only the international evidence which encompasses many education systems with widely differing institutional structures has the potential to show whether institutions have important consequences for student performance. The data used in this study combines the TIMSS international micro data base, which includes both student performance data and data on family and school background, with additional data on the education systems (Section 3).

In the microeconometric student-level investigation, many institutional features show strong and significant effects on students' performance, while different categories of resource inputs again reveal no clear positive impact (Section 4). Expenditure per student and class size actually show adverse effects, while positive effects seem to arise from a sufficient equipment with instructional materials and from teachers' education and experience. The educational background of a student's family is strongly conducive to her performance. After controlling for resource and family background effects, the influence of institutional incentive mechanisms adds substantially to an explanation of differences in student performance in mathematics and in science. Both centralized examinations and the size of the private schooling sector are shown to have statistically significant positive effects on student performance. In general, school autonomy also seems to yield positive effects. However, the effect of the distribution of responsibilities between schools and administration differs for different kinds of decisions. On the one hand, extensive decision-making powers of schools over the purchase of supplies, the hiring and rewarding of teachers, and the organization of instruction have statistically significant positive effects on student performance. On the other hand, it seems also positive for students' learning if responsibility for the determination of the curriculum syllabus, for the approval of textbook lists, and for the school budget does not lie at the school level. Once an educational task or educational funding is delivered by the administration, implementation at an intermediate level seems to be more conducive to student performance than implementation either at the central or at the local level of administration.

The effects of increased decision-making influence of teachers again depend on the relevant domain of decision-making and on the way in which it is exerted. While an increased influence of individual teachers on the curriculum has statistically significant positive effects on student performance, an increased influence of school teachers acting collectively in teacher unions has statistically significant negative effects. It is conducive to student performance if the class teacher has influence on the choice of supplies, but it seems detrimental if she has influence on the amount of subject matters to be taught. Extra time devoted by teachers to student assessment has a strong positive effect on student performance, while there is no linear relationship between minutes of homework assigned and student performance. Parents' influence on class teaching seems conducive to students' test score performance, while there remain some ambiguities concerning the overall effect of parents' involvement. All these empirical findings support the argumentation of the institutional economics applied to the education sector.

To assess the extent to which institutional differences can account for the cross-country differences in student performance levels, country-level education production functions are estimated in Section 5. They show that institutions strongly matter for cross-country differences in students' educational performance, while increased resource inputs do not contribute to increased performance. Controlling for indicators of parents' education levels and resource inputs, three indicators of institutional features of the education system have strong and statistically significant effects on country-level student performance. Increased school autonomy in supply choice and increased scrutiny of performance assessment lead to superior performance levels, and a larger influence of teacher unions in the education process leads to inferior performance levels. Together, the variables explain three quarters of the cross-country variation in mathematics test scores and 60 percent of the variation in science test scores, whereas previous studies which focused on family and resource effects explained only up to one quarter of the cross-country variation in student performance tests.

2 Institutional Economics Applied to the Education Sector

2.1 The Impact of Institutions on the Effectiveness of Resource Use

In studying the economic forces at work in the education sector, one is easily led to the simple production-function argumentation that more inputs such as smaller classes, higher teacher salaries, or more teaching material should lead to higher schooling output in the form of improved educational performance of students. However, this would require an efficient use of resources in the sense that given inputs are used in a performance-maximizing way. The reason that this is assumed in the study of production processes in other sectors of the economy is that competition in the market place forces producers to use resources efficiently because otherwise they would lose out against their competitors. It is the incentives elicited by prices and competition that create the efficient input-output link.

Generally, this incentive mechanism is not at work in the public education sector. For that reason, we cannot simply presuppose that the educational input-output relation is efficient. Instead, we have to look at the institutional structures that prevail in the schooling system and at the monetary and intrinsic incentives they create for the different groups involved in educational production. As Landsburg (1993, p. 3) put it, "Most of economics can be summarized in four words: 'People respond to incentives.'" Therefore, to understand the economic forces at work in the education sector, I analyze the incentives influencing the different agents involved in the production of education and the different institutional structures which create these incentives.

Institutions are the constraints devised by humans which constitute the rules of the game in a society, thereby structuring human interactions (North 1994). Institutions enclose formal and informal rules and their enforcement instruments. Within the education system, relevant institutions include the ownership structure of schools, the rules governing examinations, or the formally and informally determined distribution of decision-making powers between the different agents involved. The set of given institutions creates a system of property rights, i.e. rights of agents to use resources. That is, institutions determine who is eligible to make decisions on the use of resources in different areas. In addition, institutions determine the provision of information and the rewards and penalties which the agents get in response to their actions (Furubotn and Richter 1997). Thereby, institutions define and limit the set of choices of individuals and form the incentive structure of a society.

While institutions are the rules of the game, the agents who are the players in this game act within this system of rules. I assume that individual agents act rationally, i.e. they maximize their objective functions subject to the constraints set by institutions. Therefore, they respond to the incentives created by the set of given institutions. The behavior of the agents involved in educational production is reflected in their decisions on the allocation of resources across different functional categories (e.g., number of teachers, teachers' salaries, instructional material) and on the effectiveness of the use of these resources. This in turn affects the outcome of the education process, namely the performance of the students.

Consequently, institutions influence student performance by creating a system of rights to decide on resource allocation which establishes the incentives that steer agents' behavior in a particular direction. In North's (1994, p. 359) explanation of economic performance, "[i]nstitutions form the incentive structure of a society, and the political and economic institutions, in consequence, are the underlying determinants of economic performance." In the same consequence, the political and educational institutions are the underlying determinants of educational performance.

Institutions are not per se created in a way that ensures efficiency. Quite to the contrary, in the education system, there are a lot of problems of agency, incomplete contracts, and adverse incentives which work against an efficient use of resources. The institutions governing the education process can be viewed as a network of principal-agent relationships. Within these relationships, a principal has an (explicit or implicit) contract with an agent to act on his behalf. The agent is self-interested, and he enjoys some informational advantage over the principal (asymmetric information). The self-interest of the agent might conflict with the principal's interest, and the informational advantage will make it costly (or even impossible) for the principal to monitor the actions of the agent completely. This leads to adverse incentives, giving the agent some leeway to act opportunistically - i.e. in his own interest instead of the principal's interest - without being penalized. While it might be in the interest of the "ultimate" principal in the education process (say, the parents) to maximize student performance with given resources, the interests of the different agents will lead to a misallocation across different inputs and an inefficient use of the inputs.

A (still hugely simplifying) picture of the network of principal-agent relationships in educational production looks as follows: Voters (including parents) entrust the government with the task of ensuring education for the children. The government hands the implementation over to the administration. The administration transfers the task of schooling provision to school management (usually exercised by heads of school or school governing boards). And school management employs teachers and teaching aides for tuition of the children. Each of these contracts is laden with problems of monitoring. There is no clear-cut property right of students or parents to decide how the money for their education is spent. Instead, all the agents involved respond to the incentives set by the institutions: They can use the room created by imperfectly monitored contracts to advance their own interest. They can divert resources from the use of maximizing educational outcome to the use of advancing their own objectives.

It would be an over-simplification of reality to assume that the different groups of agents maximize a single objective each. In reality, each group of agents faces multifarious interests, and the institutional structure can change the relative costs and benefits of advancing one objective or the other. While teachers clearly have a genuine interest in increasing their income at a given work-load or decreasing their work-load at a given income, no one will deny that most teachers also get pleasure from seeing their students learning much and thus raise their welfare level. Furthermore, teachers might face negative consequences from their heads of school or from parents when they are doing a bad job. Thus, teachers often face conflicting interests, and their relative advancement may be easier or harder in different institutional surroundings. If the performance of students is observed, the achievement of higher student performance will have a higher pay-off for teachers than if it is not. Likewise, if teachers have a lot of leeway to decrease their work-load without facing negative consequences, this will have adverse effects on student performance relative to a situation where they have less leeway.

Parents are probably the actors with the clearest unidimensional interest in a high level of their children's performance. While the students themselves certainly have an interest in their own performance, they will weigh this objective against other objectives such as the amount of leisure time and the possibility of making and losing friends through studying less or more. In the same way as with teachers and students, the school management and the administration will face a trade-off between advancing students' performance and reducing their work-load, while they also care for their own monetary payoff and for their school's or district's reputation. Finally, in the public-choice view, the government's interest lies in its re-election, so that it will do whatever it has to do to increase the likelihood of its re-election. This in turn will be influenced by the ability of the different interest groups to lobby for their objectives.[2]

The advancement of their own interests by the different groups of agents may lead to two kinds of inefficiencies in the allocation and use of schooling resources. First, it may be in the interest of some agents to make inefficient use of given resources (although resources may be allocated efficiently across different inputs). E.g., a teacher may be inclined to use part of a lesson for more pleasant things than stressfully teaching mathematics, as long as this lack of mathematical tuition is not monitored. Second, the agents' interest may lead to a misallocation of resources across functional categories (causing inefficiency even if these resources were then used effectively). If it is in the interest of a group of agents with decision power over resource allocation to over-spend on one input relative to others, the marginal productivity of this input would be lower than that of the other inputs, leading to a student performance level inferior to a situation of efficient spending. E.g., if teachers have a say in budgetary matters, they may want to increase spending on teachers at the expense of spending on instructional material, so that the marginal product of material inputs is higher than the marginal product of teacher inputs and schooling output could be higher at the given expenditure level.

Therefore, "there is an enormous gap between children sitting in a classroom and an increase in human capital" (Pritchett and Filmer 1999, p. 223), reflected in the quality of education. An increase in expenditure per student does not necessarily have to lead to increased student performance. Likewise, lower class sizes do not necessarily have to go hand in hand with better educational outcome. The classes may already be so small that the marginal productivity of a reduction in class size (an increase in teachers per

2 Given that there is a huge number of teachers in many parliaments in the world, the potential of teachers to lobby for their objectives might be substantial.

student) is negligible. Even more, the input "teacher per student" may not be used with the same effectiveness everywhere. If a more productive way of using resources in bigger classes outweighs any potential positive effect of smaller classes, there could even be an adverse effect of class size.

In the following, a more detailed analysis is given of the different kinds of educational institutions, the incentives they create for different educational agents, and probable consequences for students' performance. The institutions analyzed relate both to the structure of the schooling system at large and to the decision-making powers and incentives of teachers, students, and parents.

2.2 Institutional Features Related to Schools and Administration

The structure of the institutional system within education determines who has the power to decide on which task. Four main institutional features of the education system at large and their possible impact on student performance are dealt with here: centralized examinations; the distribution of decision-making power between schools and administration; the distribution of decision-making power between different levels of administration; and the extent of competition from private educational institutions in the system.

Central Examinations

Of the 39 education systems analyzed in this study, 15 have some kind of centralized examinations in the sense that a central decision-making authority has exclusive responsibility for or gives final approval of the content of examinations. The institution of centrally and thus externally set examinations profoundly alters the incentive structure within the education system. Central exams signal the achievement of a student relative to an external standard, thereby making students' performance comparable to the performance of students in other classes and schools.[3] This simplifies the monitoring of the performance of students, teachers, and schools.

With centrally set examinations, students get marks relative to the country mean, so that the performance of students is made observable and transparent. Furthermore, it is observable whether the bad performance of a given student is an exception within a class or whether the whole class taught by one teacher is doing badly relative to the country mean. Therefore, parents (and students) have the information they need to initiate action because they can observe whether the teacher (and/or the student) is accountable for the bad performance. If, by contrast, students get marks relative to the class mean only, the performance of the class relative to the country mean will be unobservable and parents will have no information to intervene. As a consequence of the institutional setting, the agents' incentives are fundamentally altered. Given central examinations, the leeway of the teachers to act opportunistically is reduced and the incentives to use resources more effectively are increased. Through central examinations, agents are made accountable to their principals: parents can assess the performance of their children, of the teachers, and of the schools; the head of a school can assess the performance of her teachers; and the government and administration can assess the performance of different schools. Thus, a strong case can be made for a positive link between centralized examinations and student performance on the basis of incomplete contracts and monitoring in the education system.

Furthermore, central examinations change the students' incentive structure (Bishop 1997). By creating comparability to an external standard, they improve the signaling of academic performance to advanced educational institutions and to potential employers, so that students' rewards

---

[3] For a more detailed description of the characteristics of "curriculum-based external exit examination systems" see Bishop (1997, 1999).

for learning should grow and become more visible. This should increase students' incentives to perform well, by increasing and making better use of their own resources spent on education (their time and attention). In addition, grading relative to class performance gives students an incentive to lower average class performance because this allows the students to receive the same grades at less effort. The cooperative solution of students to maximize their joint welfare is for everybody not to study very hard. Thus, with grades relative to the class level, students have an incentive to distract teachers from teaching a high standard and to apply peer pressure on other students in the class not to be too studious (Bishop 1999). With centralized external examinations, in contrast, these incentives are no longer given because inferior class work will only harm the students. Thus, central examinations should have a positive effect on student performance also through the channel of changed incentives of students.

Distribution of Responsibilities between Schools and Administration

A second institutional feature of the education system is the division of decision-making authority between administration and schools. For example, schools have a very high degree of decision-making autonomy in the Netherlands, while they do not have much autonomy in Greece, Norway, or Portugal.[4] On the one hand, increased decision-making power at the school level establishes freedom to decide within schools, which is a pre-requisite for competition and for the possibility to respond to demands from parents. The actors within the schools should have the decentralized knowledge to choose the best way of teaching for their students (if they have incentives to do so), a kind of knowledge probably not given at the

---

[4] As measured by the OECD indicator on the distribution of decision-making responsibilities in the education systems, less than 25 percent of educational decisions are taken at the school level in Greece, Norway, and Portugal, while the Netherlands have the highest degree of school autonomy with 73 percent of decisions taken at the school level (see Section 3.3 for more information on this indicator).

administrative level. On the other hand, increased decision-making power at the school level increases the schools' leeway to act opportunistically, unless decisions can be fully monitored and the extent to which educational objectives are met can be fully evaluated. Instead of leading to incentives to increase student performance, decentralized decision-making might thus lead to the advancement of adverse incentives as long as accountability for the increased decision-making power cannot be ensured. Consequently, external standard setting and control would be needed to restrict the opportunistic leeway of schools.

These two tendencies invoked by an increase in the decision-making autonomy of schools - better use of decentralized knowledge and increased scope for opportunism - should have opposite impacts on the effectiveness of resource use and thereby on student performance. Which direction of impact is the superior one depends on the area of decision-making. There are decisions where centralization (decreased school autonomy) is likely to have positive net effects on student performance, and there are decisions where it is likely to have negative net effects.

If decisions on standard setting and performance control are centralized, a lowering in a school's tuition standards will become easily transparent to parents and administration. This helps in the monitoring of schools' actions, thereby changing the schools' incentives against a misuse of resources. Through a centralized basic curriculum, the amount of what schools should teach is fixed and cannot easily be watered down by the interests of the agents at the school level as long as an external performance control is in place. Furthermore, knowledge on what students should be taught and on how their achievement should be measured may be equivalent or even superior at the central level relative to the school level. Therefore, centralized standard and control decisions should have positive net effects on student performance.[5] Likewise, centralized decisions on the size of the school budget should benefit the overall effectiveness of resource use and thus student performance, since agents at the school level have huge adverse incentives when it comes to the amount of resources available. It is clearly in the self-interest of decision-makers at the school level to collect additional funds for themselves or for resources which lighten their work-load.

In contrast, knowledge on which process or personnel-management decision is favorable to students' learning will be superior at the school level. Heads of school have better knowledge than the administration on which tuition structures are best for their schools, which teacher deserves a pay rise or a promotion, and which teacher is the right one to hire for the school. Likewise, individual teachers should be best in choosing the right textbooks and other kinds of supplies and in the organization of instruction. School autonomy should increase the effectiveness of resource use in these decision areas. Furthermore, school autonomy in process or personnel decisions does not generate much leeway to act opportunistically because hiring bad teachers or choosing bad textbooks is not in the interest of school personnel. Therefore, decentralization of process and personnel decisions should have positive net effects on student performance, given that standards, performance, and budgetary levels are centrally controlled.

Distribution of Responsibilities between Administrative Levels

In the preceding paragraphs, it was argued as if "the administration" was one single body. In reality, there are administrative authorities at the local, regional, state, and central levels in many countries. In the United States, for example, most of educational decision-making and basically all fund allocation takes place at the local level of government, while in Germany,

---

5 Additionally, as shown by Costrell (1994), a centralized system of standard-setting will result in higher educational standards than a decentralized system because decentralization reduces a district's marginal benefit of a higher standard and raises its marginal cost.

the responsibility for planning and purchasing educational resources is mainly with the intermediate level and in Greece, most educational decisions and basically all funding take place at the central level. Thus, the division of responsibilities for funds and decision-making between local, intermediate, and national authorities establishes another feature of the institutional system of education which may influence the educational outcome. Once responsibility lies with the administration, the question is which level should take over the tasks so that the highest effectiveness of resource use is achieved.

Again, different effects run counter to one another. At the local level, more decentralized knowledge is available and the administration is more directly accountable to parents. However, the administration will also have much closer ties with the school personnel, increasing the possibilities for successful lobbying of school-based interest groups and for collusion. Local administrators and school personnel might collude on the determination of the level and use of funds, so that an opportunistic resource allocation ensues. The central administrative level is more remote from the agents within the school. On the one hand, this should make collusion harder to achieve. On the other hand, monitoring of actions and resource use from the central level is elusive because of information problems (Hoxby 1999).[6] In addition, a self-interested central administration will find it easier to develop excessive bureaucracy and to divert resources at the central level.

Since both the local and the central level of administrative decision-making face serious deficiencies, an intermediate level might be better positioned to run educational administration. An intermediate level of administration is too far away from schools for serious local lobbying and

---

6 Hoxby (1999) emphasizes the benefits of decentralized Tiebout residential choices as a solution to the information problem. However, her model does not consider political-economy effects of lobbyism and collusion, and she concedes that there may be serious flaws in the Tiebout process.

collusion, but it is possibly superior to the central level in monitoring schools. Ultimately, it is an empirical question which administrative level performs best.

Private Schools

In general, production of basic education is run publicly all over the world. However, different education systems show differing degrees of private provision of schooling. For example, three quarters of Dutch students attend schools which are managed privately. Japan and Korea are the countries with the largest shares of privately managed schools which are also financially independent of public funding. At the other extreme, countries such as Australia, Austria, the Czech Republic, Denmark, France, Germany, Hungary, Iceland, Norway, Spain, and Sweden have virtually no financially independent private schools. When private schools are available, parents with the aim of increasing their children's educational performance can choose whether to send them to a particular private school. Through the institution of private ownership, the head of a private school has a clear monetary incentive to use resources efficiently so as to maximize student performance, because this would make parents choose her school. Therefore, she will try to improve the monitoring of her teachers. Furthermore, private provision circumvents many monitoring problems within governmental and administrative entities. While private as opposed to public provision of education cannot eliminate all the monitoring problems inherent in the education process, private schools may thus nevertheless decrease the number of difficult-to-monitor principal-agent relationships and face greater incentives to tackle the remaining ones.

By giving parents additional choice, private educational institutions introduce competition into the public education system. Because the loss of students to private institutions may have adverse consequences for the heads of public schools, increased competition from private schools should also have a positive effect on the effectiveness of resource use in nearby public schools. Thus, private ownership of property rights and competition generally establish incentives that work in the direction of efficient resource use. In a similar way, Shleifer (1998) shows that from a contracting perspective, private ownership of schools, combined with choice and competition, establishes strong incentives for cost reduction and qualitative innovation which are missing in publicly run schools. Chubb and Moe (1990) argue that public schools tend to be overbureaucratized and ineffective because they are governed by institutions of direct democratic control, while private schools tend to possess autonomy and the characteristics of an effective organization because they are governed by markets.

The different degrees of competition from private educational institutions in schooling systems across countries should therefore be a cause of differences in student performance. A larger share of privately managed educational institutions should go hand in hand with superior student performance.

2.3 Institutional Features Directly Involving Teachers, Students, and Parents

As described before, educational output is produced in a multi-layer principal-agent system. The section on the structure of the education system at large treated institutional settings at the level of schools and the administration. The responsibilities and incentives of three further groups of agents remain to be analyzed: teachers, students, and parents.

Teachers' Influence

Teachers are probably the most important external determinants of students' learning. They are agents in a contract to teach students. Within this contract, they have a lot of freedom on how to pursue their teaching, since many of their actions cannot be monitored. If teachers get a lot of influence on expenditure allocation in the education sector, they will use it to promote their own interests (Pritchett and Filmer 1999). The main interests of teachers are in their own financial well-being and small work-load on the one hand and in the achievement of their students on the other hand. The institutional setting determines teachers' incentives to behave either conducive or detrimental to student performance. Explicit or implicit rewards and penalties will tilt their behavior in one direction or the other.

There are two dimensions in which teachers' influence can be analyzed. First, the effect of teacher influence may differ between teachers acting individually and teachers acting collectively through a teacher union. The very aim of teacher unions is to promote the interests of teachers, and to defend them against the interests of other interest groups.[7] Therefore, they will focus on the interests which are not advanced by the other interest groups. The main interests of teachers which are not advanced by others are to increase their pay and to decrease their work-load. Furthermore, teacher unions can exert collective bargaining power - as opposed to individual teachers and to other groups of agents which can less easily be organized -, and they will advance the interest of the median teacher, which favors a leveling out of salary scales instead of merit differentiation. Thus, other things being equal, a high degree of decision-making power of teacher unions should favor behavior detrimental to student performance. By contrast, when teachers act individually, the benefits of an increased use of their decentralized knowledge at the classroom level stand against their interest to decrease their work-load and may even outweigh them. Thus, the effect of increased decision-making power of individual teachers on student performance is ambiguous.

---

[7] Hoxby (1996) stresses that teacher unions have both the interest to obtain more generous inputs and the potential to lower the effectiveness of input use.

Second, the effect of teacher influence may differ between different decision-making areas. Similar to the argument before for the distribution of responsibility between schools and administration, a high degree of teacher influence on process decisions, such as what supplies to be bought or which textbooks to be used, should be conducive to student performance, because teachers are the agents who know best how to teach their students and because there is not much leeway to exploit this kind of decision-making power opportunistically. By contrast, a high degree of teacher influence in determining teacher salary levels or on decisions which are related to their work-load, such as the amount of subject matters to be taught, will be detrimental to student performance, because this creates huge incentives for teachers to behave selfishly.

Students' Incentives

All efforts to increase the amount of educational resources and the effectiveness of their use will probably be in vain unless the student who shall reap the educational benefits is open to learning and has incentives to study. Learning requires the time and active engagement of students. It therefore stands in competition for students' time with other, presumably more pleasant uses. The incentives to study - rewards that increase the benefits of studying and penalties that increase the cost of failing to do so - will determine the intensity of a student's investment in learning (Bishop 1999).

These incentives are in turn set by the institutional framework. Central examinations increase the benefits of studying and the costs of not studying. Likewise, the scrutiny with which achievement is observed and marked by teachers determines the extent to which studying is rewarded and laziness penalized. Increased intrinsic and extrinsic rewards for learning make it more worthwhile to study. An increased assignment and monitoring of homework should increase student effort and learning as well, tilting the trade-off between studying and other uses of student's time in favor of studying. Both of these institutional features should have positive effects on student performance.

Parents' Influence

Parents are the only actors within education who have a relatively undisturbed interest in the educational performance of their children. They have a clear interest in schooling resources being used efficiently. Therefore, increased decision-making and monitoring powers of parents should tilt the incentives of educators in favor of a more effective use of resources and a superior educational outcome - at least as long as parents view education mainly as an investment in their children's human capital and not as consumption.

Parents' participation in the educational process is limited by the opportunity cost of their time. Institutions which give parents a greater say enhance the benefits of participation and make parental involvement more likely. As a result, an institutional setting which ensures increased overall participation of parents in the educational process and gives parents greater influence on decisions on what to be taught should be beneficial for students' performance.

3 Data and Methodological Concepts

The argumentations of institutional economics applied to the education sector in the preceding section have yielded several hypotheses on the link between institutions and student performance which can be tested by the estimation of a microeconometric education production function. As the latest, largest, and most extensive international student achievement test ever conducted, the Third International Mathematics and Science Study (TIMSS) provides both student performance data and student, teacher, and school background data for representative samples of students in about 40 countries. Section 3.1 describes TIMSS and motivates the use of international evidence. Section 3.2 discusses the use of TIMSS test scores as the dependent variable, while Section 3.3 deals with the explanatory variables, comprising data on the students' family background, school resources, and institutional structures retrieved both from TIMSS and from other sources.

3.1 The TIMSS International Micro Data Base

In order to show that institutions have important consequences for student performance, one has to show that institutional variation leads to a variation in student performance. But there is usually no significant variation in many institutional features within a single country on which an enlightening analysis could be based (Chubb and Moe 1990).[8] However, there are big differences across countries in such institutional features as the size of the private schooling sector, the centralization of examinations and of other decision-making powers, and the responsibilities and influence of different educational agents. Therefore, to test the institutional hypotheses, I use international evidence.

Until now, the only evidence available on the effects of institutionalized incentive regimes like decentralized management on educational performance is based on case studies of experiments and specific programs (Hanushek et al. 1994). Econometric investigation has so far not used the huge international evidence that exists, presumably shying away from analyzing educational processes across different countries and cultures. The perspective taken in this paper is that economic principles influence the

[8] This led Chubb and Moe (1990) to base their empirical analysis on a comparison of the public and the private schooling sector in the United States. This empirical approach and their choice of key concepts and analytical models has been heavily criticized in the literature; see e.g. Bryk and Lee (1992).

actions of human beings in any country or culture. People respond to incentives everywhere. Therefore, taking care of the economic and institutional differences which exist between countries, the international data can be used to analyze determinants within the education process.

In TIMSS, extensive efforts have been made to deal with the challenges associated with comparing achievement across countries, cultures, and languages through careful planning, cooperation among the participating countries, standardized procedures, and rigorous attention to quality control. TIMSS was conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA), which has gathered 40 years of experience with international comparative studies on educational achievement and learning contexts. In 1994-95, TIMSS tested representative samples of students in more than 40 countries.[9] In addition to mathematics and science achievement scores, the TIMSS international data base contains a myriad of educational variables representing background information about teaching and learning collected from students, teachers, and heads of school. Altogether, TIMSS tested and gathered contextual data for more than half a million students and administered questionnaires to teachers and heads of school in 15,000 schools, thereby offering an unprecedented opportunity to examine determinants of student performance.[10]

---

9 Of the 45 countries participating in TIMSS, three (Argentina, Indonesia, and Italy) were unable to complete the steps necessary to appear in the data base. Mexico chose not to release its results. For three countries (Bulgaria, Philippines, and South Africa), no background data are included in the international data base because of insufficient quality. Since in Belgium, the Flemish and French education systems participated separately, data files for 39 education systems are available: Australia, Austria, Canada, Colombia, Cyprus, Czech Republic, Denmark, England, Flemish Belgium, France, French Belgium, Germany, Greece, Hong Kong, Hungary, Iceland, Iran, Ireland, Israel, Japan, Korea, Kuwait, Latvia, Lithuania, Netherlands, New Zealand, Norway, Portugal, Romania, Russian Federation, Scotland, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Thailand, and United States.

10 For more information about the design, development, implementation, and analyses of TIMSS, see the internet homepage at http://timss.bc.edu/.

Countries participating in the study were required to administer tests to the students in the middle school years, but could choose whether or not to participate at the primary and final school years. Therefore, this paper focuses on the middle school years, where students enrolled in the two adjacent grades containing the largest proportion of 13-year-old students were tested, which are seventh- and eighth-grade students in most countries. This micro data set includes data on more than 260,000 individual students which form a representative sample of a population of more than 30 million students in 39 education systems.

The TIMSS achievement tests used the experience gained by the predecessor studies. They were developed through an international consensus-building process involving input from international experts in mathematics, science, and measurement, and were endorsed by all participating countries. Based on a curriculum framework developed by educators from around the world, test specifications were developed that included items representing a wide range of mathematics and science topics and eliciting a range of skills from students. The TIMSS tests include items requiring students both to select the appropriate response, to provide a short answer to a question or problem, and to provide a more elaborate response or explanation.

In addition to testing students, TIMSS collected contextual information about instruction and learning through student, teacher, and school questionnaires. The data base includes information on approximately 1,500 instructional, school, and home background variables. The students who participated in TIMSS completed questionnaires about their demographics, home background, and classroom and out-of-school activities. The mathematics and science teachers of sampled students responded to questions about their professional training and education, their instructional practices, their responsibilities for decision-making in several areas, and about class sizes and the availability of materials. The heads of schools provided information on school characteristics and resources, the degree of centralization of decision-making, the allocation of responsibility for different tasks, and topics like the extent of parents' participation. Additionally, some system-level information was provided by the national research coordinators of each participating country.

Basic results of TIMSS have been published in international achievement reports (e.g., Beaton et al. 1996a, b). They have usually shown that while home factors are strongly related to mathematics and science achievement, the relationship is less clear between achievement and various instructional variables. However, these reports contain mainly uni-variate, within-country analyses.

3.2 Test Scores in Mathematics and Science as Measures of Educational Performance

While test scores of cognitive achievement in mathematics and science may be reasonable measures of the output in central areas of schooling and may thus capture important aspects of the human capital of students, they certainly do not reflect the whole array of socially and economically valuable human capital. First, there are many problems with constructing meaningful and internationally comparable standardized tests of cognitive skills. Second, there are many other subjects in school apart from mathematics and science - many of which do not easily lend themselves to standardized achievement tests. And third, there are many valuable, mostly non-cognitive skills formed outside schools, mainly in families and later in firms.

However, many arguments warrant an analysis based on the test scores of cognitive achievement in mathematics and science. First of all, much care was taken with the TIMSS test design, so that it used probably the best technique available up to now to measure cognitive skills in mathematics and science. TIMSS tried to remedy shortcomings of earlier approaches. The study implemented rigorous procedures to prevent bias, to ensure comparability in school and student sampling, and to assure quality in test design and development, data collection, scoring procedures, and analysis. TIMSS covered a wide range of topics and capabilities in the two subjects and elicited a range of skills from the students. In mathematics, the content areas included fractions and number sense, geometry, algebra, measurement, proportionality, and data representation, analysis, and probability. In science, they included earth science, life science, physics, chemistry, and environmental issues and the nature of science. Many different kinds of performances were expected of students, encompassing categories such as understanding simple information, performing routine procedures, using complex procedures, solving problems, proving, communicating, and investigating the natural world.

In their assessment of the relative merits of standardized multiple-choice tests and authentic open-ended-response tests, Hanushek et al. (1994, p. 137) point out that both have (often offsetting) advantages and disadvantages and thus recommend the utilization of a combination of both. This is exactly the strategy used by TIMSS. Approximately one-fourth of the questions (designed to represent approximately one-third of students' response time) were in the free-response format, requiring students to generate and write their own answers. Some free-response questions asked for short answers while others required extended responses where students needed to show their work. The remaining questions used a multiple-choice format.

Combining the performance in the different questions of a subject, proficiency was mapped onto an international scale with a mean of 500 and a standard deviation of 100, yielding the international achievement scores. TIMSS was designed to ensure international comparability. The curriculum framework on which the TIMSS achievement test is based takes care that the achievement items are appropriate for the students of all participating countries and reflect their current curriculum.[11] All in all, the TIMSS test results are most probably the best one can currently get in measuring student achievement in mathematics and science.

In a modern economy where economic success is increasingly attributed to advances in research and development, high levels of cognitive skills in mathematics and science are important to provide for able engineers and scientists. At the aggregate, cross-country level, Hanushek and Kim (1999) have shown that educational quality as measured by comparative tests in mathematics and science has a consistent, stable, and strong influence on economic growth. Furthermore, many studies have shown that there are large returns to higher achievement test scores in mathematics and science in the labor market (Boissiere et al. 1985, Bishop 1989, 1992). These earnings advantages to higher achievement on standardized mathematics tests have grown recently (Murnane et al. 1995). For students dropping out of school, basic cognitive skills are also an important determinant of later earnings (Tyler et al. 1999). Earnings in the labor market are the most reasonable measure of the economically valuable human capital of a person, encompassing the returns to several dimensions of human capabilities. Therefore, the fact that measures of cognitive skills in mathematics and science are a good predictor of future earnings suggests that this measure can actually serve as a - possibly weak - proxy for other skills.

Finally, it is widely accepted that it is much harder to quantify and measure performance in other subjects and non-cognitive skills than it is to measure performance in mathematics and science. Consequently, it is

[11] A test-curriculum matching analysis conducted by TIMSS showed that omitting those items for each country which measure topics not addressed in the curriculum had little effect on the overall pattern of achievement results across all countries.

sensible to confine the application of econometric methods to the production of those skills which are readily quantifiable, keeping in mind the limitations which may be implied.

Therefore, the test scores should be taken for what they are good: They give a reasonable measure of the cognitive achievement of students in mathematics and in science. An analysis of the production of these cognitive skills is of interest in itself as they constitute an important part of socially and economically valuable human capital. Furthermore, there is evidence suggesting that implications derived from the analysis of these cognitive skills expand well into other parts of human capital.

3.3 Data on Student Background, Resources, and Institutions

A complete list of the variables used in this study and their descriptive statistics is given in Table 1. The sources of the variables - as reported in the column labeled "Origin" - are TIMSS student, teacher, and school background questionnaires, as well as country-level data on the education systems obtained also from OECD educational indicators. For dummy variables, the column labeled "True" reports the percent of students for which the state expressed by the dummy is true. For discrete variables, the international means and standard deviations are reported.

Data from TIMSS Student, Teacher, and School Questionnaires

Student-level information (marked "St" in the column "Origin" of Table 1) is used to control for students' background and family characteristics. To capture the fact that performance should differ between the two adjacent grades in which students were tested, the dummy "upper grade" is set equal to one for each student in the upper one of the two grades tested. In addition, two countries (Sweden and Switzerland) have tested students in a third grade above the other two, which is captured by the dummy "above upper grade". Data on the students' age and sex are included, as well as dummies showing whether the student was born in the country where she goes to school, whether she lives with both parents, and whether at least one parent was born in the country. Furthermore, four dummies are included representing the highest educational level achieved by the students' parents, as well as four dummies for the number of books in the students' home, which acts as a proxy for the educational and social background of the family. In the production functions to be estimated, the coefficients on the dummies for parents' education show the performance of students with parents who achieved some secondary education, finished secondary education, had some education beyond the secondary level, and finished university relative to students with parents who only had primary education. Likewise, the dummies for books at home show the performance of students with four ranges of numbers of books (11-25 or enough to fill one shelf, 26-100 or enough to fill one bookcase, 101-200 or enough to fill two bookcases, and more than 200 or enough to fill three or more bookcases) relative to students with none or very few (up to 10) books at home.

Table 1: Descriptive Statistics

Variable	Origin[1]	True (Percent)	Mean	Std. Dev.	Imputed (Percent)	Countries Missing
Test scores						
Mathematics score	St		507.6	99.3	0.0	-
Science score	St		506.8	98.4	0.0	-
Student and family characteristics						
Upper grade	St	51.3			0.0	-
Above upper grade	St	1.8			0.0	-
Age (years)	St		13.8	0.9	1.4	-
Sex (female)	St	50.2			0.4	-
Born in country	St	94.1			7.6	FRA JPN
Living with both parents	St	87.3			7.3	JPN
Parent born in country	St	91.9			7.7	JPN
Parents' education	St				13.2	GBR JPN
Some secondary		14.2				
Finished secondary		31.8				
Some after secondary		24.4				
Finished university		18.1				
Books at home	St				6.0	JPN
11-25		14.8				
26-100		33.0				
101-200		19.4				
More than 200		24.9				
Community location					10.2	KWT
Geographically isolated area	Sc	3.0				
Close to the center of a town	Sc	39.7				
GDP per capita (intl. $)	C		15404.0	6993.1	0.0	-
Resources and teacher characteristics						
Expenditure per student (intl. $)	C		3242.8	1941.7	0.0	-
Mathematics class size (no. of stud.)	T		28.4	10.7	26.8	-
Science class size (no. of students)	T		28.8	10.2	35.8	-
Student-teacher ratio	Sc		25.7	61.2	21.0	GBR
No shortage of materials	Sc	41.0			12.3	SCO
Great shortage of materials	Sc	12.5			12.3	SCO
Instruction time (minutes per year)	Sc		47226.5	8248.6	30.0	GRC JPN KWT
Mathematics teacher characteristics						
Teacher's sex (female)	T	54.1			12.0	-
Teacher's age (years)	T		40.8	9.3	12.0	-
Teacher's experience (years)	T		16.2	9.4	13.2	-
Teacher's education	T				22.2	AUS BFL/R DNK JPN
Secondary		18.8				
BA or equivalent		64.8				
MA/PhD		15.3				
Science teacher characteristics						
Teacher's sex (female)	T	52.4			15.0	-
Teacher's age (years)	T		40.7	9.2	15.0	-
Teacher's experience (years)	T		15.7	9.3	16.3	-

(to be continued)

Table 1 (continued)

Variable	Origin[1]	True (Percent)	Mean	Std. Dev.	Imputed (Percent)	Countries Missing
Teacher's education	T				24.9	AUS BFL/R DNK JPN
Secondary		15.9				
BA or equivalent		66.6				
MA/PhD		16.7				
Institutional settings						
Central examinations						
Central examinations	C	34.2				-
External exams influence curriculum	Sc	15.2			20.0	BFL/R GBR KWT NOR SCO
Distribution of responsibilities between schools and administration						
Central curriculum	C	77.2				-
Central textbook approval	C	51.3				-
School responsibility						
School budget	Sc	90.5			20.4	GBR NOR SCO
Purchasing supplies	Sc	96.2			21.3	GBR NOR SCO
Hiring teachers	Sc	73.1			17.4	GBR NOR SCO
Determining teacher salaries	Sc	27.9			20.3	GBR NOR SCO BF/R
Teachers' influence						
Teachers' responsibility						
School budget	Sc	3.2			20.4	GBR NOR SCO
Purchasing supplies	Sc	10.0			21.3	GBR NOR SCO
Hiring teachers	Sc	0.1			17.4	GBR NOR SCO
Determining teacher salaries	Sc	0.1			20.3	GBR NOR SCO BFL/R
Strong influence on curriculum						
Teacher individually	Sc	24.3			17.5	BFR GBR KWT NOR SCO
Subject teachers	Sc	48.9			18.0	BFR GBR KWT NOR SCO
School teachers collectively	Sc	45.0			17.4	BFR GBR KWT NOR SCO
Teacher unions	Sc	1.7			17.9	BFR GBR KWT NOR SCO
Math. teacher has strong influence on						
Money for supplies	T	4.9			13.7	GBR
Kind of supplies	T	13.2			13.7	GBR
Subject matter	T	28.5			13.3	GBR
Textbook	T	19.9			13.6	GBR
Sci. teacher has strong influence on						
Money for supplies	T	7.4			17.0	GBR ISR
Kind of supplies	T	20.1			16.8	GBR ISR
Subject matter	T	40.6			16.5	GBR ISR
Textbook	T	24.8			16.9	GBR ISR
Students' incentives						
Mathematics						
Scrutiny of exams (hours p. w.)	T		2.5	1.6	14.1	DNK
Homework (minutes per week)	T		99.3	80.0	22.4	-
Science						
Scrutiny of exams (hours p. w.)	T		2.4	1.6	16.9	DNK
Homework (minutes per week)	T		38.8	40.9	32.9	AUT

(to be continued)

Table 1 (continued)

Variable	Origin[1]	True (Percent)	Mean	Std. Dev.	Imputed (Percent)	Countries Missing
Parents' influence						
Parents influence curriculum	Sc	1.8			17.1	BFR GBR KWT NOR SCO
Mathematics						
Uninterested parents limit teaching	T	8.3			25.9	JPN
Interested parents limit teaching	T	4.8			24.2	FRA
Parent-teacher meetings (hours p. w.)	T		0.6	0.6	15.6	DNK
Science						
Uninterested parents limit teaching	T	6.4			34.3	JPN
Interested parents limit teaching	T	4.2			32.5	FRA
Parent-teacher meetings (hours p. w.)	T		0.6	0.6	18.8	DNK
(in percent:) *Distribution of responsibilities between schools and administration*						**Number of Countries Available**
School autonomy	CO		16.5	11.3		21
School level decisions						
Overall	CO		40.1	16.9		21
Organization of instruction	CO		80.1	13.1		21
Personnel management	CO		29.9	29.0		21
Planning and structures	CO		24.2	18.3		21
Resources	CO		26.2	19.4		21
Distribution of responsibilities between administrative levels						
Central government decisions						
Overall	CO		25.2	21.5		21
Organization of instruction	CO		9.9	11.1		21
Personnel management	CO		31.8	32.7		21
Planning and structures	CO		41.0	30.6		21
Resources	CO		18.2	28.9		21
Government level of final funds						
Funds provided at local level	CO		38.0	36.4		22
Funds provided at central level	CO		28.1	34.4		22
Private schools						
Private enrollment	CO		13.8	14.0		23
Independent private enrollment	CO		4.3	6.9		23
Public exp. on private institutions	CO		10.6	17.2		26
Public exp. on indep. private inst.	CO		0.4	0.9		26

Statistics are unweighted.

[1] St = student, T = class teacher, Sc = school, C = country, CO = OECD country data.

Data obtained from TIMSS teacher and school questionnaires are used to show effects of school resources and institutional features. To do so, the student-specific data on achievement test scores and student characteristics has been merged with the class-level teacher data and with the school-level data. Teacher data are available separately for mathematics and science teachers. If a student had more than one teacher in either mathematics or science, the teacher who taught the most minutes per week to that student was chosen when merging student and teacher data.

The teacher data (T) give class-level information on class size and the class teacher's characteristics, behavior, and influence. Teacher characteristics considered are the teacher's sex, age, years of teaching experience, and education. The teacher's education is again given in dummy form, showing the effect of the completion of secondary education, bachelor degree, and masters or doctorate degree relative to teachers without completing secondary education. The relevant data on institutional features which are used to deal with the hypotheses of Sections 2.2 and 2.3 is derived from qualitative survey data. Teachers were asked how much influence they had on the amount of money to be spent on supplies, on what supplies are purchased, on the subject matter to be taught, and on specific textbooks to be used. The four-item response possibilities of the questionnaire ranged from "none" to "a lot". To be able to assess the impact of teachers' influence in the different decision-making areas on students' performance, dummy variables were created from these qualitative data which are set equal to one for each teacher who answered that she had "a lot" of influence in the respective field. In the same way, a dummy variable was created signifying that parents uninterested in their children's learning and progress limit "a great deal" how the teacher teaches his class, as well as one for limitations by parents interested in their children's learning. Further data gives information on the scrutiny of exams as measured by the time (in hours per week) which the teacher spends outside the formal school day on preparing or grading exams, on minutes of homework per week assigned by the teacher, and on the time which the teacher spends outside the formal school day on parent-teacher meetings.

The school-level information (Sc) was given by the principals of the schools tested. It includes dummies for schools located in geographically isolated communities and for schools located close to the center of a town (the residual category being schools located in villages or on the outskirts of a town), as well as dummies for schools whose principal thinks that the capacity to provide instruction is "not" or "a lot" affected by the shortage or inadequacy of instructional materials like textbooks. Data is given on the ratio of total student enrollment to professional staff at the school, where the latter includes principals, assistant principals, department heads, classroom teachers, teacher aides, laboratory technicians, and learning specialists. An instruction-time variable combines information on the duration of a typical instructional period in minutes, on the number of instructional periods of an average school day, and on the number of instructional days in the school year, each reported separately for the lower and the upper grade. Apart from these resource-related data, qualitative assessment of the institutional setting is given. Principals were asked who, with regard to their school, had primary responsibility for different activities, including formulating the school budget, purchasing supplies, hiring teachers, and determining teacher salaries. Possible answers were "teachers," "department heads," "the principal," "the school's governing board," and "not a school responsibility." Dummies were created for each activity showing whether primary responsibility was with the school and whether it was with teachers. Finally, six dummies show whether external examinations/standardized tests, each teacher individually, teachers of the same subject as a group, teachers collectively for the school, teacher unions, and parents had "a lot" of influence in determining the curriculum that is taught in the school.

Data Imputation

For the TIMSS questionnaire variables, missing data was a problem. While some students, teachers, and school principals failed to answer individual questions, some other questions were not at all administered in some countries. Since dropping students with missing data on some explanatory variables from the regression analysis would severely reduce the sample size, delete the information available on the other explanatory variables, and introduce sample selection bias, imputation of missing values was chosen.[12] Values were imputed using the data of those students with non-missing values on the variable of interest and data on a set of "fundamental" explanatory variables available for all students.

For each student *i* with missing data on a specific variable *V*, a set of "fundamental" explanatory variables *F* with data available for all students was used to impute the missing data in the following way. Let *S* denote the set of students with available data for *V*. Using the students in *S*, the variable *V* was regressed on *F*. For *V* being a discrete variable, ordinary least squares estimation was used for the regression. For *V* being a dichotomous (binary) variable, the probit model was used. If *V* was originally (before deriving dummies) a polychotomous qualitative variable with multiple categories, an ordered probit model was estimated. The coefficients from these regressions and the data *F*(*i*) were then used to impute the value of *V*(*i*) for the students with missing data. For the probit models, the estimated coefficients were used to forecast the probability of occurrence associated with each category for the students with missing data, and the category with the highest

[12] Data on individual students for whom non-imputed data were available for less than 25 variables of the standard regression specification of Table 2 were dropped from the sample. This excluded 614 students from the original sample of 267,159.

probability was imputed. The set of "fundamental" explanatory variables *F* included: the student's sex; the student's age; two dummies on the grade level which the student attended; four dummies on the parents' education level; four dummies on the number of books in the student's home; three dummies on the type of community in which the school is located; gross domestic product (GDP) per capita of the country; and educational expenditure per student in secondary education in the country. The small amount of missing data within *F* was imputed by taking the average value at the lowest level available, that is, class average, school average, or country average. For the three countries which did not administer one of the variables in *F*, these missing data were imputed by the method outlined above, using the remaining variables in *F* for the regression. The percentage of imputed values for each variable is given in Table 1, as well as the names of the countries which did not provide internationally comparable data for each variable. Results of robustness tests against dropping from the sample observations with imputed data for each individual variable are reported in Section 4.4.

Country-Level Data

At the country level (C), TIMSS reported the degree of centralization regarding decision-making about examinations, curriculum syllabi, and textbooks (Beaton et al. 1996a, pp. 17-19). Dummies report whether the central level of decision-making authority within the (national or regional) education system had exclusive responsibility for or gave final approval of the content of examinations, the syllabi for courses of study, and the list of approved textbooks. As a measure of the country's level of development which sets the general background for the educational opportunities of the students, data on GDP per capita (in current international dollars, 1994) is included, obtained from World Bank (1999) statistics. Unfortunately, the TIMSS data base does not contain expenditure data. In lack of expenditure data at the local level, current public expenditure per student in secondary education at the national level in 1994 was calculated on the basis of data from UNESCO Statistical Yearbooks, converted into international dollars with purchasing-power-parity exchange rates from the World Bank (1999).

Further country-level data on institutional features of the education system at large - concerning the level of decision-making and private involvement - were obtained from the 1997 and 1998 volumes of the OECD educational indicators (CO), which are devised to provide a quantitative description of the comparative functioning of education systems. Most of these indicators were gathered in the UNESCO/OECD World Education Indicators (WEI) program, which includes a number of non-OECD-member countries. The number of countries for which each indicator is available is reported in Table 1.

With regard to the organization of the schooling system, the OECD has gathered information on the distribution of decision-making responsibilities among key shareholders of education. These agents are central, state, provincial, sub-regional, and local authorities as well as schools.[13] The indicator relates to public lower secondary education and is available for the school year 1997/98 only.[14] It is based on 35 decision items included in a survey completed by panels of national experts and refers to actual decision-making practice (not to non-binding formal regulations). The variables used in this paper represent the percentages of educational decisions taken at the school level and at the central government level (OECD 1998, Tables E5.1 and E5.2). The general indicator (termed "overall" in Table 1) is calculated to give equal importance to each of the following four decision-making domains: organization of instruction, personnel management, planning and

---

[13] Information on names and numbers of decision-making units per decision-making level for each country can be found in Annex 3 of OECD (1998).

[14] This should not be a problem, however, since organizational structures in the education system are relatively stable.

structures, and resource allocation and use. Data is also available for each of the four domains separately. Decisions on the organization of instruction include the determination of school attendance, promotion, repetition, and teaching and assessment methods. Decisions on personnel management comprise the hiring, dismissal, career influence, and duties of staff as well as the fixing of salary scales. As for planning and structures, decisions such as the creation and closure of schools and grades, the design of programs of study, and the setting of qualifying examinations for a certificate are combined. Resource decisions include the allocation and use of resources for different kinds of expenditure. Finally, the variable "school autonomy" captures what percentage of decisions are taken at the school level in full autonomy, as opposed to after consultations with or within frameworks from other educational authorities (OECD 1998, Table E5.3).

Another OECD educational indicator reports which share of final public funds is spent by the different levels of government (after transfers between the different government levels), showing whether local, intermediate, or central authorities are the final purchasers of educational resources (OECD 1998, Table B6.1a). The variables on the "government level of final funds" report the respective shares of total public funds provided at the local level and at the central level, giving information on the division of responsibility for and control over the funding of education between local, intermediate, and national authorities.

Data on the share of private enrollment in total enrollment refer to the school year 1994/95 (derived from OECD 1997, Table C1.1a). Educational institutions are classified as either public or private according to whether a public agency or a private entity has the ultimate power to make decisions concerning the institution's affairs (ultimate management control). That is, public institutions are institutions controlled and managed directly by a public education authority or agency or controlled by a body whose members are appointed by a public authority, while private institutions are controlled and managed by a non-governmental organization (e.g. a church or a business enterprise). As a sub-group of the private institutions, government-independent private institutions are those private institutions which receive less than 50 percent of its core funding from government agencies, while government-dependent private institutions primarily depend on funding from government sources for their basic educational services. The variable "public expenditure on (independent) private institutions" contains the share of public educational expenditure going to (independent) private educational institutions in the financial year 1995 (OECD 1998, Table B6.2).

4 Microeconometric Results

4.1 Estimating Micro Education Production Functions

To determine the influence of student background, resources, and institutions on students' educational performance, an education production function can be estimated of the form

$$t_i = B_i \alpha + R_i \beta + I_i \gamma + \varepsilon_i$$

where t is the test score of student i, B are the measures of the student's background, R are the measures of resource use, I are the measures of institutional features surrounding the student's learning (R and I are measured at the classroom, school, and country level), ε is an error term, and α, β, and γ are the parameters to be estimated. The γ parameters can be used to test the hypotheses of the institutional economics in the education sector.

Studies such as Lee and Barro (1997) and Hanushek and Kim (1999) have used country-level data to analyze the determination of students' performance. These macro education production functions cannot control for individual influences on a student's performance. Apart from the impossibility of properly controlling for other influences, investigations at the country or provincial level are restrained to the analysis of system-level institutional determinants like central examinations (as performed by Bishop 1997, 1999), but they cannot deal with institutional features working at lower levels. Bishop (1999, p. 395) concedes that, "important as [curriculum-based external exit examination systems] may be, they are not the only or even the most important determinant of achievement levels." To assess other institutional determinants of students' achievement, one has to look below the country level. Hence, the relevant level at which to perform the analysis is the individual student (not the class, school, district, or country), because this directly links a student's performance to her teaching environment. The estimation of such a microeconometric education production function provides the opportunity to control for individual background influences on student performance when looking at the influence of resources and institutions, to assess the influence of the relevant resource and teacher characteristics with which a student is faced, and to look at the institutional features relevant to the individual student.

In using student-level data, attention has to be given to the complex data structure given by the survey design and the multi-level nature of the explanatory variables. As is common in educational survey data, the TIMSS sampling design includes varying sampling probabilities for different students as well as stratified and clustered data (see Martin and Kelly 1998), giving rise to the three econometric problems of sampling weights, data stratification, and data clustering. As regards sampling weights, the TIMSS procedure was designed to achieve nationally representative student samples by stratified sampling within each country. To avoid bias in the estimated equation and to obtain nationally representative coefficient estimates from stratified survey data, weighted least squares (WLS) estimation using stratum weights has to be employed so that the proportional contribution of each stratum in the sample to the parameter estimates is the same as would have been obtained in a complete census enumeration (Klein 1974, pp. 409-412). DuMouchel and Duncan (1983) show that the use of a WLS estimator is especially relevant in an omitted-predictor model, which is certainly given in the estimation of an education production function where the innate ability of each student remains unmeasured. Therefore, the following WLS regressions use weights which assure that each student is weighted according to her probability of selection so as to yield representative samples within each country and to give each country the same weight in the international estimation.

As regards data stratification, each country was sampled separately, so that sampling was done independently across countries, fixing the division into countries in advance. In consequence, the TIMSS data is stratified by country. As regards data clustering, the TIMSS sampling procedure had a two-stage clustered sample design within each country, with the first stage yielding a sample of schools and the second stage yielding a sample of classrooms (Gonzalez and Smith 1997). Thus, the primary sampling unit (PSU) in TIMSS was the school. Individual students who go to the same school may share some characteristics which are not perfectly captured by the included observable variables. Furthermore, the data set is characterized by a hierarchical data structure with data collected at different levels. As the resource and institutional variables are not measured at the student level but at the classroom level or the school level (see above), the observations on these variables for students who share the same class or school depend on one another. As a result, observations in the same PSU are not independent, so that the structure of the error term in the equation given above may be more complicated than conventional econometric methods assume. The problem with the conventional formulas for the computation of standard errors which ignore the cluster design of the data is that they overstate precision by ignoring the dependence of observations within the same PSU. To accommodate the potentially non-usual error structure of the hierarchically structured data, a method is required for correcting the estimated standard errors of the least squares regression. One such method, usually referred to as robust linear regression (RLS), combines the WLS regression with robust estimates of standard errors which recognize the stratification and clustering of the survey design.[15] RLS relaxes the independence assumption and requires only that the observations be independent across the PSUs, allowing any amount of correlation within the PSUs. Thus, RLS allows one to estimate appropriate standard errors when many cases share the same value on some but not all independent variables.[16]

In the following tables with the estimated parameters of education production functions, robust standard errors based on RLS are presented in addition to conventional (raw) standard errors. The robust standard errors are based on countries as strata and schools as PSUs. As the highest level of clustering, schools were chosen as PSUs, thereby allowing any degree of dependence within schools. Therefore, the reported robust standard errors are actually upper bounds for the coefficients of those explanatory variables which are measured at the student or classroom level. The marks signaling

---

15 For a more detailed description of RLS, see Deaton (1997, pp. 74-78).

16 A related method often used in educational research to account for the hierarchical data structure are hierarchical linear models (HLM), also known as multi-level models (cf. Bryk and Raudenbush 1992, Goldstein 1999). Cohen and Baldi (1998) show that RLS is superior to HLM because it does not require HLM's assumptions of random and normally distributed effects. When these assumptions are violated by outliers or by a skewed error distribution, HLM significantly underestimates the standard errors of higher-level parameters and gives biased estimates, respectively, while RLS provides estimates which are both consistent and robust. Only when the assumptions are met that errors at all levels are truly normally distributed, estimates of lower-level coefficients are slightly more efficient under HLM than RLS. Thus, while RLS may sacrifice some efficiency at the lower level relative to HLM and thus lead to overly conservative estimates, HLM can lead to invalid inference under moderate violations of its assumptions.

significance levels in the results tables are based on these robust variance estimates.

4.2 Family Background and Resource Effects

Table 2 shows OLS and WLS regression results for the mathematics achievement score. The results apply to a sample of 266,545 students in the middle school years from 39 schooling systems. While the results do not differ considerably between the OLS and the WLS estimation, the following discussion refers to the WLS estimates. Furthermore, significance statements are based on the robust variance estimation which accounts for the clustered data structure.

Student and Family Characteristics

Before being able to test the hypotheses of the institutional economics in the education sector, effects of differences in student characteristics and school resources have to be controlled for. Students in higher grades perform considerably better than students in lower grades. In mathematics, students in 8th grade scored 40.3 points above students in 7th grade (holding all other influences constant), and 9th-grade students scored 100.3 points above 7th-grade students. After controlling for these differences in grade levels, the age of students is negatively related to performance, probably reflecting a grade repetition effect. After controlling for the other influences, girls performed 7.6 points lower than boys. Students being born in the country in which they attend school, students living with both parents, and students who had at least one parent born in the country where they attend school performed better than otherwise.

The educational level achieved by the students' parents was strongly positively related to the students' educational performance. Relative to students whose parents only had primary education, students with parents finishing secondary or higher education performed considerably better.

Parents' attending of some secondary schooling without finishing it did not make a sizable difference relative to only having primary education. The effect captured by the variable "books at home," which proxies for the educational and social background of the family, was even stronger than that of the highest educational level achieved by the parents. The performance level increases steadily from students having less than 10 books at home over less than 25, 100, and 200 books to more than 200 books. Students of schools located in geographically isolated communities performed worse than students from more urban areas. Finally, as a control for the overall level of development of the country in which the student lives, GDP per capita is positively related to mathematics achievement. All these effects of student and family characteristics are statistically highly significant.

Student and family background effects on science achievement, reported in Table 3, are very similar to the case of mathematics achievement. While being qualitatively identical, the quantitative effect differs to some extent for some variables. For example, the lead of boys' performance over girls' performance was 8.5 points larger in science than in mathematics.

Table 2: Microeconometric results for mathematics performance

Dependent variable: TIMSS international mathematics test score. Standard errors in parentheses.

	OLS		WLS			
	Coeff.	Raw S.E.	Coeff.	Raw S.E.	Robust S.E.	Std. Coeff.
Constant	426.985	(4.360)	482.793*	(4.211)	(13.916)	
Student and family characteristics						
Upper grade	38.773	(0.425)	40.342*	(0.424)	(1.086)	0.202
Above upper grade	99.486	(1.464)	100.313*	(1.513)	(3.906)	0.127
Age	-9.884	(0.244)	-14.183*	(0.231)	(0.779)	-0.135
Sex	-7.229	(0.343)	-7.634*	(0.346)	(0.878)	-0.038
Born in country	8.372	(0.813)	9.199*	(0.816)	(1.338)	0.021
Living with both parents	15.276	(0.514)	12.099*	(0.519)	(0.814)	0.040
Parent born in country	5.132	(0.715)	3.983†	(0.722)	(1.602)	0.011
Parents' education						
Some secondary	0.069	(0.707)	-3.989*	(0.702)	(1.553)	-0.014
Finished secondary	25.755	(0.654)	26.475*	(0.660)	(1.454)	0.123
Some after secondary	12.046	(0.695)	15.130*	(0.700)	(1.515)	0.066
Finished university	36.600	(0.734)	39.724*	(0.746)	(1.619)	0.152
Books at home						
11-25	10.999	(0.755)	10.326*	(0.749)	(1.360)	0.037
26-100	37.317	(0.705)	35.846*	(0.701)	(1.444)	0.168
101-200	47.570	(0.761)	46.713*	(0.756)	(1.543)	0.186
More than 200	55.145	(0.753)	54.269*	(0.750)	(1.562)	0.235
Community location						
Geographically isolated area	-14.707	(1.040)	-18.502*	(1.085)	(3.385)	-0.030
Close to the center of a town	2.451	(0.361)	1.598	(0.363)	(1.479)	0.008
GDP per capita	0.004	(5.9e-5)	0.004*	(5.8e-5)	(2.1e-4)	0.240
Resources and teacher characteristics						
Expenditure per student	-0.009	(2.1e-4)	-0.006*	(2.1e-4)	(6.9e-4)	-0.106
Class size	0.912	(0.018)	1.176*	(0.019)	(0.090)	0.122
Student-teacher ratio	0.011	(0.003)	0.006	(0.003)	(0.007)	0.004
No shortage of materials	8.525	(0.387)	7.230*	(0.394)	(1.585)	0.036
Great shortage of materials	-1.480	(0.563)	-5.925†	(0.554)	(2.393)	-0.020
Instruction time	3.7e-4	(2.3e-5)	3.1e-4*	(2.3e-5)	(8.4e-5)	0.025
Teacher characteristics						
Teacher's sex	5.634	(0.372)	5.727*	(0.374)	(1.345)	0.029
Teacher's age	-0.712	(0.033)	-0.667*	(0.033)	(0.124)	-0.062
Teacher's experience	1.075	(0.032)	1.038*	(0.033)	(0.121)	0.097
Teacher's education						
Secondary	11.151	(1.674)	15.682*	(1.569)	(5.206)	0.062
BA or equivalent	10.919	(1.648)	10.571†	(1.542)	(5.105)	0.050
MA/PhD	20.860	(1.694)	25.576*	(1.596)	(5.411)	0.090

(to be continued)

Table 2 (continued)

	OLS Coeff.	Raw S.E.	WLS Coeff.	Raw S.E.	Robust S.E.	Std. Coeff.
Institutional settings						
Central examinations						
Central examinations	17.842	(0.434)	16.062*	(0.402)	(1.435)	0.045
External exams influence curriculum	10.740	(0.539)	4.271‡	(0.524)	(2.199)	0.016
Distribution of responsibilities between schools and administration						
Central curriculum	15.585	(0.539)	10.776*	(0.519)	(1.783)	0.048
Central textbook approval	10.053	(0.474)	9.559*	(0.460)	(1.563)	0.078
School responsibility						
School budget	-5.362	(0.663)	-5.852†	(0.683)	(2.450)	-0.017
Purchasing supplies	-2.288	(0.976)	0.538	(0.997)	(3.488)	0.001
Hiring teachers	13.959	(0.454)	12.723*	(0.471)	(1.772)	0.055
Determining teacher salaries	6.539	(0.455)	10.588*	(0.464)	(2.112)	0.046
Teachers' influence						
Teachers' responsibility						
School budget	-15.478	(1.032)	-13.318*	(1.100)	(3.805)	-0.022
Purchasing supplies	11.361	(0.602)	14.148*	(0.642)	(2.576)	0.040
Hiring teachers	-4.317	(5.413)	-10.294	(6.197)	(21.456)	-0.003
Determining teacher salaries	-16.874	(5.153)	-11.069	(5.492)	(20.995)	-0.003
Strong influence on curriculum						
Teacher individually	9.709	(0.442)	11.952*	(0.446)	(1.730)	0.051
Subject teachers	-2.980	(0.473)	-6.855*	(0.476)	(1.897)	-0.034
School teachers collectively	-9.333	(0.459)	-12.659*	(0.459)	(1.836)	-0.063
Teacher unions	-27.532	(1.367)	-32.329*	(1.370)	(5.979)	-0.042
Class teacher has strong influence on						
Money for supplies	2.800	(0.905)	-0.815	(0.909)	(3.734)	-0.002
Kind of supplies	-2.701	(0.593)	-0.627	(0.606)	(1.997)	-0.002
Subject matter	-0.613	(0.414)	-0.830	(0.420)	(1.585)	-0.004
Textbook	-0.322	(0.480)	2.687	(0.478)	(1.913)	0.011
Students' incentives						
Scrutiny of exams	4.410	(0.109)	4.749*	(0.110)	(0.429)	0.078
Homework	-0.006	(0.002)	0.001	(0.002)	(0.010)	0.001
Parents' influence						
Parents influence curriculum	-0.949	(1.314)	3.714	(1.390)	(5.516)	0.005
Uninterested parents limit teaching	-12.546	(0.672)	-10.107*	(0.656)	(2.756)	-0.029
Interested parents limit teaching	-8.879	(0.871)	-10.860*	(0.825)	(4.090)	-0.025
Parent-teacher meetings	-5.966	(0.277)	-6.152*	(0.283)	(1.021)	-0.039
Observations	266545		266545			
Schools (PSUs)	6107		6107			
Countries	39		39			
R^2 (adj.)	0.22		0.22			

* Significant at the 1 percent level based on robust standard errors.
† Significant at the 5 percent level based on robust standard errors.
‡ Significant at the 10 percent level based on robust standard errors.

Resources and Teacher Characteristics

The estimated effects of the amount of resources used on student performance are consistent with most of the literature in that no strong positive relationship exists between spending and student performance (see, e.g., Hanushek 1996). In fact, instead of resulting in higher student performance, higher educational expenditure per student (measured at the country level) and smaller class sizes (measured at the classroom level) are statistically significantly related to inferior mathematics and science results.[17] The same holds for smaller ratios of students to total professional staff at the school, which is statistically insignificantly different from zero. For all three of these resource variables, the observed effects show an adverse direction.

In contrast to the measured effects of teaching staff, the equipment with facilities has the expected effect when measured by the subjective assessment of the principals of the schools. Students in schools whose principals reported that the capacity to provide instruction is not affected by the shortage or inadequacy of instructional materials scored 7.2 points higher in mathematics relative to students in schools with a little or some limitation (6.5 in science), while students in schools with great shortage of materials scored 5.9 (11.6) points worse.

Instruction time (in minutes per year) at the relevant grade level of the school is statistically significantly positively related to student performance in mathematics and science. While the relative importance of the explanatory dummy variables can be directly evaluated on the basis of their regression coefficients (the coefficient of dummies reports the conditional test score difference between students with and without the characteristic of interest), standardized coefficients (also reported in Tables 2 and 3 for the

WLS estimation) can be used to compare the relative importance of the discrete explanatory variables. For example, a change of 1 standard deviation in instruction time is related to a change of only 0.025 standard deviations in the mathematics test score, while a 1 standard-deviation change in the mathematics class size is related to a change of 0.122 standard deviations in mathematics performance.

As for teacher characteristics, students of female teachers score statistically significantly higher than students of male teachers in both mathematics and science. Controlling for the teacher's age, more years of experience are positively related to students' performance. Conversely, controlling for the teacher's experience, the teacher's age is negatively related to students' performance. This may reflect positive effects of teaching experience in combination with negative effects of age differences between teachers and students, presumably due to increasing difficulties of intergenerational understanding and declining motivation of aging teachers. Teachers' level of education is positively related to students' performance. Relative to students of teachers who did not complete secondary education, students of teachers who finished secondary education scored considerably better. While the effect was largest for teachers with a masters or doctorate degree (25.6 points in mathematics and 32.1 points in science), teachers with a bachelor degree or equivalent (10.6 and 12.4) did not achieve the same results for their students as teachers who just completed secondary education plus perhaps some teacher training (15.7 and 24.2). Overall, the effects of teachers' education levels were larger in science than in mathematics.

---

[17] When calculating robust standard errors on the basis of countries as PSUs, taking account of the fact that expenditure per student is measured at the country level, the expenditure effect turns statistically insignificant in mathematics.

Table 3: Microeconometric results for science performance

Dependent variable: TIMSS international science test score. Standard errors in parentheses.

	OLS Coeff.	Raw S.E.	WLS Coeff.	Raw S.E.	Robust S.E.	Std. Coeff.
Constant	409.230	(4.525)	455.626*	(4.315)	(11.881)	
Student and family characteristics						
Upper grade	43.897	(0.434)	46.568*	(0.433)	(0.990)	0.235
Above upper grade	99.908	(1.491)	105.354*	(1.544)	(3.536)	0.134
Age	-6.128	(0.249)	-10.116*	(0.236)	(0.708)	-0.097
Sex	-15.546	(0.349)	-16.130*	(0.352)	(0.753)	-0.081
Born in country	10.428	(0.828)	11.195*	(0.831)	(1.305)	0.026
Living with both parents	9.320	(0.524)	7.437*	(0.529)	(0.800)	0.025
Parent born in country	13.686	(0.728)	12.536*	(0.736)	(1.400)	0.034
Parents' education						
Some secondary	-2.142	(0.720)	-5.226*	(0.715)	(1.469)	-0.019
Finished secondary	17.830	(0.667)	20.067*	(0.674)	(1.284)	0.094
Some after secondary	8.421	(0.708)	10.423*	(0.714)	(1.330)	0.046
Finished university	30.827	(0.747)	34.304*	(0.760)	(1.424)	0.132
Books at home						
11-25	12.381	(0.769)	12.251*	(0.763)	(1.153)	0.044
26-100	35.629	(0.718)	34.174*	(0.715)	(1.248)	0.161
101-200	50.483	(0.775)	48.862*	(0.770)	(1.348)	0.196
More than 200	59.954	(0.767)	57.494*	(0.764)	(1.370)	0.250
Community location						
Geographically isolated area	-4.163	(1.058)	-7.371†	(1.106)	(3.397)	-0.012
Close to the center of a town	-0.958	(0.369)	-2.215‡	(0.371)	(1.306)	-0.011
GDP per capita	0.004	(6.0e-5)	0.004*	(5.9e-5)	(2.0e-4)	0.264
Resources and teacher characteristics						
Expenditure per student	-0.011	(2.2e-4)	-0.010*	(2.2e-4)	(6.4e-4)	-0.186
Class size	0.362	(0.019)	0.477*	(0.020)	(0.060)	0.047
Student-teacher ratio	0.010	(0.003)	0.009	(0.003)	(0.007)	0.006
No shortage of materials	6.998	(0.394)	6.543*	(0.402)	(1.374)	0.032
Great shortage of materials	-7.375	(0.573)	-11.595*	(0.565)	(2.138)	-0.039
Instruction time	3.4e-4	(2.3e-5)	3.0e-4*	(2.4e-5)	(6.8e-5)	0.024
Teacher characteristics						
Teacher's sex	5.947	(0.377)	7.801*	(0.378)	(1.166)	0.039
Teacher's age	-0.272	(0.033)	-0.216‡	(0.033)	(0.113)	-0.020
Teacher's experience	0.457	(0.033)	0.445*	(0.033)	(0.115)	0.041
Teacher's education						
Secondary	19.882	(2.032)	24.243*	(1.801)	(4.940)	0.091
BA or equivalent	11.241	(1.993)	12.378†	(1.758)	(4.859)	0.059
MA/PhD	25.575	(2.034)	32.106*	(1.806)	(5.042)	0.119

(to be continued)

Table 3 (continued)

	OLS Coeff.	Raw S.E.	WLS Coeff.	Raw S.E.	Robust S.E.	Std. Coeff.
Institutional settings						
Central examinations						
Central examinations	8.598	(0.437)	10.650*	(0.405)	(1.302)	0.024
External exams influence curriculum	2.329	(0.550)	-4.364†	(0.536)	(1.881)	-0.016
Distribution of responsibilities between schools and administration						
Central curriculum	5.319	(0.552)	5.573*	(0.530)	(1.649)	0.031
Central textbook approval	5.563	(0.474)	6.157*	(0.460)	(1.346)	0.052
School responsibility						
School budget	-2.065	(0.674)	-3.451	(0.695)	(2.356)	-0.010
Purchasing supplies	1.939	(0.996)	2.867	(1.016)	(3.308)	0.006
Hiring teachers	6.235	(0.461)	5.247*	(0.478)	(1.473)	0.023
Determining teacher salaries	11.381	(0.462)	15.162*	(0.473)	(1.817)	0.067
Teachers' influence						
Teachers' responsibility						
School budget	-9.172	(1.048)	-4.583	(1.116)	(3.025)	-0.008
Purchasing supplies	7.052	(0.613)	6.837*	(0.653)	(2.062)	0.019
Hiring teachers	6.817	(5.518)	7.595	(6.315)	(6.002)	0.002
Determining teacher salaries	-9.640	(5.249)	-6.048	(5.600)	(16.342)	-0.002
Strong influence on curriculum						
Teacher individually	8.711	(0.450)	10.768*	(0.455)	(1.536)	0.046
Subject teachers	-2.129	(0.481)	-4.573*	(0.485)	(1.625)	-0.023
School teachers collectively	-3.084	(0.468)	-5.034*	(0.468)	(1.575)	-0.025
Teacher unions	-18.901	(1.393)	-18.395*	(1.395)	(5.533)	-0.024
Class teacher has strong influence on						
Money for supplies	3.764	(0.764)	6.876*	(0.791)	(2.255)	0.018
Kind of supplies	3.871	(0.516)	4.566*	(0.530)	(1.520)	0.018
Subject matter	-0.429	(0.382)	-1.213	(0.380)	(1.186)	-0.006
Textbook	-0.978	(0.453)	-1.016	(0.459)	(1.379)	-0.004
Students' incentives						
Scrutiny of exams	0.513	(0.116)	0.444	(0.117)	(0.406)	0.007
Homework	-0.031	(0.004)	-0.043*	(0.005)	(0.014)	-0.017
Parents' influence						
Parents influence curriculum	0.264	(1.339)	5.041	(1.416)	(4.411)	0.006
Uninterested parents limit teaching	-12.980	(0.776)	-11.003*	(0.758)	(2.649)	-0.028
Interested parents limit teaching	-0.295	(0.952)	-1.333	(0.922)	(3.394)	-0.003
Parent-teacher meetings	-2.293	(0.289)	-2.662*	(0.290)	(0.859)	-0.017
Observations	266545		266545			
Schools (PSUs)	6107		6107			
Countries	39		39			
R^2 (adj.)	0.18		0.19			

* Significant at the 1 percent level based on robust standard errors.
† Significant at the 5 percent level based on robust standard errors.
‡ Significant at the 10 percent level based on robust standard errors.

In sum, the relationship between school resources and student performance is ambiguous. Expenditure per student and class size show adverse effects, while equipment with instructional materials and teachers' experience and education show positive effects. What is clear is that there certainly is no strong and systematic relationship between resource use and student performance. In light of the argumentation of Section 2, this is actually not surprising: Within the institutional setting of public schooling systems, the incentives of the agents involved do not clearly point in the direction of increased student performance. So the next section investigates whether differences in the incentive structures determined by the institutional features of the education systems have significant effects on student performance.

4.3 Institutional Effects

Schools and Administration I

Central Examinations

Central examinations should have a positive impact on the efficiency of resource use in the education system and on the performance of students. In accordance with this hypothesis, students in countries with centralized examination systems scored 16.1 points higher in mathematics and 10.7 points higher in science than students in countries without centralized examinations.[18] The evidence suggests that central examinations matter in explaining international differences in the productivity of schooling systems.

[18] With only 39 independent observations, the effect of this country-level variable entails some degree of uncertainty. Calculating robust standard errors with countries as PSUs leaves the effect of central examinations in mathematics statistically significant only at the 15 percent level, while it turns statistically insignificant in science. When increasing the threshold of non-imputed data to a sample size of 255,018 (see Section 4.4), the mathematics effect is significant at the 10 percent level based on robust standard errors with countries as PSUs.

Hence the micro evidence corroborates the findings at the country and provincial level of Bishop (1997, 1999).

Furthermore, students in schools where external examinations or standardized tests had a lot of influence in determining the curriculum had test scores 4.3 points higher in mathematics than students in schools where this was not the case. In science, the effect is negative when imputed observations are included, while it is positive but insignificant when the observations with imputed data on this variable are dropped from the sample (see Section 4.4). The weaker effect of standardized tests in science than in mathematics may reflect that science tests lend themselves less readily to standardization.

Distribution of Responsibilities between Schools and Administration

The responsibility for decisions in several areas of the education system is distributed differently between administration and schools across countries. The question is whether schools have freedom to decide or whether there is control from above. With respect to centralized standard setting and control decisions, two country-level dummies report whether decision-making responsibilities for the syllabi for courses of study ("central curriculum") and for the list of approved textbooks ("central textbook approval") are centralized. Students in countries both with centralized curriculum setting and centralized textbook approval score higher in mathematics and science than students in countries without these decisions being centralized.[19] However, the absolute size of both these effects is smaller than the effect of centralized examinations.

The division of decision-making authority between administration and schools is also relevant in financial and process decisions. These have been

19 With robust standard errors calculated on the basis of countries as PSUs - taking account of the fact that there are only 39 independent observations on these two variables -, these effects cannot be statistically significantly distinguished from zero, however.

measured at the school level, and the category of school responsibility encompasses the decision-making powers of teachers, department heads, principals, and schools' governing boards. Students in schools which had primary responsibility for formulating the school budget had lower scores in mathematics (5.9 points) and science (3.5 points) than students in schools which did not primarily determine their own budget.[20] That is, taking away responsibility for the amount of resources available from the school level is conducive to student performance. By contrast, school autonomy in process decisions on purchasing supplies goes hand in hand with superior achievement of students.[21] This is also true for decisions on hiring teachers, where students in schools which had freedom to decide on the hiring of teachers performed statistically significantly better in mathematics (12.7 points) and science (5.2 points) than students in schools without primary responsibility in the hiring of teachers. Likewise, students in schools which could determine teacher salaries themselves scored 10.6 points higher in mathematics (15.2 points higher in science) than students in schools without decision autonomy in this regard. Thus, school autonomy in personnel management seems highly conducive to student performance, and centralization of these decision-making powers robs schools of the opportunity to make decisions which favor their students.

In sum, the evidence supports the hypothesis that the distribution of responsibilities between schools and administration matters for the efficiency of resource use in an education system and for the educational performance of students. On the one hand, centralized decisions on standard setting, performance control, and the size of the school budget help to assure

[20] While the mathematics effect is statistically significant, the science effect is not.

[21] The low level of significance of the "purchasing supplies" coefficient in the WLS estimation for science is due to multicollinearity. Running the same regression without the other "responsibility" variables, the coefficient on "purchasing supplies" is 5.813 and statistically significant at the 10 percent level (with a robust standard error of 3.180).

that the producers of education look for the performance of students. On the other hand, school autonomy (decentralized decisions at the school level) seems to be the best way to guarantee high student performance in process and personnel-management decisions. While centralized decision-making can add coherence in curriculum coverage and control output from the center, an easing of process and personnel-management regulations may help schools' flexibility in tailoring instruction to the different needs of students. Thus, the most conducive combination seems to be a mechanism of control from above to limit school-level opportunistic behavior combined with a high degree of freedom to decide at the school level on subjects where school-level knowledge is important.

Further evidence on the distribution of responsibilities between schools and administration, as well as on the distribution of responsibilities between different administrative levels and on competition from private schools, can be obtained on the basis of the OECD educational indicators. These will be discussed after effects of institutional features directly involving teachers, students, and parents, which are based on TIMSS questionnaire data.

Teachers, Students, and Parents

Teachers' Influence

When looking at the decision-making power of specific groups of agents, the degree of freedom of teachers to decide independently on several educational topics should impact on student performance by affecting the decision-making outcome in the education system. Correspondingly, students in schools whose principals reported that teachers had primary responsibility for the school budget scored 13.3 points worse in mathematics (4.6 points in science) than students in schools where primary responsibility for the school budget was not with teachers.[22] Conversely, students scored

---

[22] The low degree of significance of the effect in science is due to a mixture of multicollinearity and data imputation. Both if the other "responsibility" variables are

14.1 points better in mathematics (6.8 in science) if teachers had primary responsibility for purchasing supplies. These two findings reflect similar effects as those reported before for the dummies representing that the two decision-making areas were primarily a school responsibility, where the school category included department heads, the principal, and the school's governing board in addition to teachers. Decisions on the amount of money to be spent should be centralized, i.e. taken away from schools and teachers, while decisions on which specific supplies to be purchased should be decentralized to schools and teachers.[23]

With regard to teachers' influence on the curriculum that is taught in the school, a clear difference arises between teachers acting individually and teachers acting collectively. On the one hand, students in schools where each teacher individually had a lot of influence on the curriculum performed considerably better than otherwise (12.0 points in mathematics and 10.8 points in science). On the other hand, students in schools where school teachers collectively or teacher unions had a lot of influence on the curriculum performed statistically significantly worse than in the case where these groups of teachers did not have a lot of influence on the curriculum. This detrimental effect of teachers exercising a collective influence on the curriculum is strongest in the case of teacher unions (-32.3 points in mathematics and -18.4 points in science).[24]

---

dropped and if observations with non-original data on the variable of teachers' responsibility for the school budget are dropped, the effect gets statistically significant and larger in absolute terms.

23 The effects of teachers being responsible for the hiring of teachers and for the determination of teacher salaries are statistically insignificant in both mathematics and science, with negative coefficients in mathematics and opposing signs of coefficients in science. This insignificance is due to the fact that only in 4 schools with a total of less than 300 tested students (out of the total of 266,545 students), heads of school reported that teachers had primary responsibility for hiring teachers (6 schools in the case of teacher salaries).

24 This finding of a negative effect of teacher unions' influence in the education system corresponds to Hoxby's (1996) result that teacher unionization can explain how

Concerning specific influence areas of individual teachers, the results are less homogenous across subjects. Students of teachers who reported that they had a lot of influence on money for supplies and on what kind of supplies are purchased performed statistically significantly better in science, whereas the effect was statistically insignificant in the case of mathematics. Students of teachers who reported that they had a lot of influence on the subject matter to be taught performed worse in science,[25] while the effect was again statistically insignificant in mathematics. Whether the class teacher is allowed to decide on the specific textbook to be used does not seem to have a significant effect on students' performance. These findings suggest that a high degree of individual teacher influence is conducive to student performance in the case of process decisions related to the choice of supplies, while it is detrimental in the case of decisions on the choice of subject matters to be taught which determines the work-load of the teacher.

Overall, the findings on teachers' influence give a clear picture. If individual teachers can make use of their decentralized knowledge on which teaching method may be best for their students, this will help students to learn more. This conclusion is corroborated by the positive effects of individual teachers influencing the curriculum that is taught in the school and of teachers having responsibility for the purchase of supplies. However, if teachers can use their decision-making powers primarily to reduce their work-load, this will hurt students' learning opportunities. This conclusion is corroborated by the negative effects of teachers' responsibility for the school budget and for the teaching load and of teachers exerting collective power

---

schools can simultaneously have worse student performance and bigger school budgets including more generous inputs. Hoxby has shown on the basis of panel data for US school districts that teacher unions increase school inputs but reduce the productivity with which these inputs are used sufficiently to have a negative overall effect on student performance, so that their primary effect is rent seeking.

[25] When observations with imputed data on this variable are dropped from the regression, the effect turns larger and statistically significant (see Section 4.4).

over the curriculum. Especially the power wielded by teacher unions, which have the explicit purpose of furthering teachers' own interests against the interests of other agents involved in the education process, is a clear example of agents acting opportunistically to further goals different from students' performance.[26]

Students' Incentives

The incentives of students to learn should be influenced by institutional features of the education system which determine the time a student spends studying and the relative benefits of studying. As reported before, instruction time in school was positively related to students' performance. Likewise, centralized examinations - which should make students' learning efforts more visible to external observers and wipe out students' incentives to lower the average performance level of the class - were shown to have a positive impact on students' educational achievement. As another measure of the extent to which studying is rewarded and laziness penalized, the amount of time outside the formal school day spent by the class teacher each week on preparing or grading student tests or exams reflects the scrutiny with which teachers observe and mark students' achievement. This scrutiny of exams has a statistically significantly positive effect on student performance in mathematics, while the positive effect in science is not statistically significant.

The amount of homework assigned by the class teacher is another measure determining the amount of time which students spend studying. However, minutes of homework per week assigned by the class teacher is statistically insignificantly related to students' performance in mathematics

[26] The results on teachers' influence are in accordance with the evidence presented by Pritchett and Filmer (1999) that inputs which provide teachers with direct benefits are generally over-used in educational production relative to inputs which contribute only to student performance.

and negatively to students' performance in science. This may reflect that minutes of homework assigned by the teacher may be very different from minutes of homework spent by each student. Alternatively, it may reflect a non-linear, more complex relationship between minutes of homework assigned and student performance. The variable reflecting minutes of homework per week is combined from two original variables reflecting how often homework is assigned per week and how many minutes it takes for an average student to complete one homework assignment. Both in mathematics and in science, homework frequency is negatively related to student performance, while homework length is actually positively related to student performance. It seems that assigning homework less often but on a more ambitious scale each is particularly conducive to students' learning. This should be more of a question for educationists, however. In any event, there is clearly no direct positive relationship between minutes per week a student spends on homework and her test score performance.

Parents' Influence

Evidence was previously reported that parents' education and the number of books in a student's home were strongly positively related to the student's educational performance. Apart from the learning environment at home, the influence which parents exert on curricular matters and on teaching in the formal education system should also impact on students' learning opportunities. Accordingly, students in schools where parents had a lot of influence in determining the curriculum scored higher both in mathematics and in science than students in schools where parents did not strongly influence the curriculum; however, these effects are not statistically significantly different from zero.

With regard to parents' influence on teaching, the class teacher reported whether parents uninterested in their children's learning and progress strongly limited how she teaches her class, e.g. because she then could not rely on parents in scrutinizing homework. The class teacher also reported whether interested parents limited class teaching, presumably by preventing her from teaching in the way she judged most suitable. Students in classes with uninterested parents strongly limiting class teaching performed 10.1 points worse in mathematics and 11.0 points worse in science relative to students in classes where teaching was not a great deal limited by uninterested parents. When interested parents were deemed a cause of limitation, students scored 10.9 points worse in mathematics. However, this effect is very small and statistically insignificantly different from zero in science. That is, even though science teachers maintained that their teaching was greatly limited by parents being excessively interested in their children's learning, this interference did not cause inferior performance of the students.

These positive effects of parents' involvement were not replicated in a positive effect of the time parents spent on meeting with teachers. In fact, the number of hours outside the formal school day reported by the class teacher to be spent on meetings with parents each week ("parent-teacher meetings") was actually negatively related to student performance. However, this may reflect the fact that teacher have more to discuss with parents of poor students than with parents of good students, so that the time spent on parent-teacher meetings is not exogenous to the students' performance. Furthermore, the hours for parent-teacher meetings are preventing the teacher from doing other useful work like the preparation and evaluation of classes and exams.

Schools and Administration II: OECD Evidence

Additional evidence on the effects of institutional features of the education system at large, encompassing features related to administration and schools, can be obtained from the institutional measures of the OECD educational indicators. Table 4 reports the WLS mathematics results for these indicators, which are all measured in percentages within a country. Since the OECD indicators are country-level variables, the number of countries equals the number of independent observations for these effects. Consequently, the standard errors reported are robust standard errors based on countries as PSUs. To save on degrees of freedom given that the OECD variables are available only for a limited number of countries participating in TIMSS and because most of the indicators are available for a different sample of countries, each row in Table 4 reports the results for a different regression, with the number of included student observations and countries and the adjusted R^2 of the regression given at the end of each row. All results reported in Table 4 are again controlling for all the student background, resource, and other institutional variables as reported in Table 2. Table 5 reports the same results for science performance.

Distribution of Responsibilities between Schools and Administration

Evidence based on TIMSS questionnaire measures presented above showed that the distribution of responsibilities between schools and administration in different educational decision-making areas has a statistically significant impact on student performance. An OECD indicator of school autonomy reports the percentage of educational decisions in a country taken at the school level in full autonomy without consultations or preset frameworks. Decisions which are not taken at the school level are taken by the administration at the central, state, provincial, sub-regional, or local level of government. The general indicator of full school autonomy ("school autonomy" in Tables 4 and 5) - which comprises the decision-making domains of organization of instruction, personnel management, planning, and resources - is statistically significantly positively related to student performance in science, and statistically insignificantly positively in mathematics. The standardized coefficients show that if the percentage of decisions taken at the school level in full autonomy increased by 1 standard deviation (equivalent to 11.3 percentage points), students scored 0.062 standard deviations (equivalent to 6.1 test score points) higher in science.

Table 4: Microeconometric results in mathematics: OECD data

Dependent variable: TIMSS international mathematics test score. WLS regression. Each row contains the result of a separate regression. Controlling for all variables of Table 2. Robust standard errors based on countries as primary sampling units in parentheses.

	Coeff.	Robust S.E.	Std. Coeff.	Obser-vations	Coun-tries	R^2 (adj.)
Distribution of responsibilities between schools and administration						
School autonomy	0.120	(0.469)	0.015	136478	21	0.19
School level decisions:						
Overall	0.227	(0.463)	0.042	134004	21	0.20
Organization of instruction	0.891 †	(0.368)	0.131	134004	21	0.20
Personnel management	-0.043	(0.295)	-0.014	134004	21	0.20
Planning and structures	0.136	(0.502)	0.027	134004	21	0.20
Resources	-0.002	(0.320)	-4.4e-4	134004	21	0.20
Distribution of responsibilities between administrative levels						
Central government decisions:						
Overall	-0.447	(0.271)	-0.097	134004	21	0.20
Organization of instruction	-1.734 *	(0.436)	-0.203	134004	21	0.21
Personnel management	-0.234	(0.197)	-0.078	134004	21	0.20
Planning and structures	-0.114	(0.189)	-0.037	134004	21	0.20
Resources	-0.371 ‡	(0.210)	-0.108	134004	21	0.20
Government level of funds:						
Funds provided at local level	-0.410 *	(0.120)	-0.150	160615	22	0.20
Funds provided at central level	-0.346 †	(0.161)	-0.126	160615	22	0.19
Private schools						
Private enrollment	0.594 †	(0.243)	0.105	170846	23	0.19
Independent private enrollment	2.909 *	(0.824)	0.195	170846	23	0.20
Public expenditure on private institutions	0.621 *	(0.159)	0.132	185786	26	0.20
Public expenditure on independent private institutions	12.124 ‡	(6.658)	0.101	185786	26	0.20

* Significant at the 1 percent level based on robust standard errors.
† Significant at the 5 percent level based on robust standard errors.
‡ Significant at the 10 percent level based on robust standard errors.

Table 5: Microeconometric results in science: OECD data

Dependent variable: TIMSS international science test score. WLS regressions. Each row contains the result of a separate regression. Controlling for all variables of Table 3. Robust standard errors based on countries as primary sampling units in parentheses.

	Coeff.	Robust S.E.	Std. Coeff.	Obser-vations	Coun-tries	R^2 (adj.)
Distribution of responsibilities between schools and administration						
School autonomy	0.523 ‡	(0.300)	0.062	136478	21	0.19
School level decisions:						
Overall	0.779 *	(0.272)	0.142	134004	21	0.19
Organization of instruction	0.808 †	(0.358)	0.116	134004	21	0.19
Personnel management	0.429 ‡	(0.215)	0.133	134004	21	0.19
Planning and structures	0.578 †	(0.272)	0.113	134004	21	0.19
Resources	0.393	(0.253)	0.076	134004	21	0.18
Distribution of responsibilities between administrative levels						
Central government decisions:						
Overall	-0.554 †	(0.239)	-0.117	134004	21	0.19
Organization of instruction	-1.178 †	(0.462)	-0.134	134004	21	0.19
Personnel management	-0.411 †	(0.148)	-0.133	134004	21	0.19
Planning and structures	-0.153	(0.172)	-0.049	134004	21	0.18
Resources	-0.417 *	(0.112)	-0.118	134004	21	0.19
Government level of funds:						
Funds provided at local level	0.023	(0.116)	0.008	160615	22	0.17
Funds provided at central level	-0.196	(0.159)	-0.070	160615	22	0.17
Private schools						
Private enrollment	0.539 *	(0.138)	0.093	170846	23	0.17
Independent private enrollment	1.257 †	(0.522)	0.082	170846	23	0.17
Public expenditure on private institutions	0.138	(0.284)	0.029	185786	26	0.18
Public expenditure on independent private institutions	4.569	(4.149)	0.037	185786	26	0.18

* Significant at the 1 percent level based on robust standard errors.
† Significant at the 5 percent level based on robust standard errors.
‡ Significant at the 10 percent level based on robust standard errors.

Likewise, including decisions at the school level that have been taken within frameworks from or after consultation with other levels of administration ("school level decisions"), increased decision-making authority at the school level - as opposed to different administrative levels - is also conducive to student performance, with the positive coefficient being statistically significant only in science.

The variables on school responsibility are also given for the four sub-groups of decisions separately. For science performance, the coefficients on school level decisions are statistically significantly positive in the decision-making domains of organization of instruction, personnel management, and planning and structures. For example, a change of 1 standard deviation (equivalent to 13.1 percentage points) in the percentage of decisions on organization of instruction taken at the school level is related to a change of 0.113 standard deviations (equivalent to 11.2 test score points) in students' science performance. For mathematics performance, the coefficients on school level decisions in the four sub-domains are not statistically different from zero, except for the decision-making domains of organization of instruction. These results on the four sub-groups of decision-making authority at the school level are the only ones with considerable differences between mathematics and science.

The fact that the effects of school responsibility are positive for science but statistically insignificant for mathematics may indicate that mathematics lends itself much easier to standardization, so that responsibilities for mathematics teaching may be taken away more easily from the school level. Since the benefits of school-level knowledge may therefore be small in mathematics, the deficiencies of school-level opportunism could bring the net effects close to zero. However, even in mathematics these effects are subject to the controlling for the school-responsibility variables measured at the school level as reported in Table 2. While the coefficient on the OECD indicator of personnel management is small and statistically insignificant, the strong positive effects of schools' primary responsibility for hiring teachers and for determining teacher salaries render the combined effect of decentralized personnel decisions positive even in the case of mathematics.

Distribution of Responsibilities between Administrative Levels

When decision-making authority lies with the administration (as opposed to the school level), the remoteness of this authority from the school level establishes another feature of the institutional system of education. Both the level of administrative decision-making and the control over educational funding may mainly be at the local, intermediate, or central level of government. The dominant level of administrative decision-making in education differs widely across countries. For example, Portugal (69 percent) and Greece (56 percent) have large percentages of educational decisions taken at the central level of government,[27] while in Belgium and the United States, the central level of government has basically no decision-making power in schooling matters. In Belgium, most of the decisions (61 percent) are taken at the lower regional level, and in the United States, the majority of decisions (69 percent) are taken at the local level. The distribution of educational funding (final purchasing of educational resources) between administrative levels follows a similar pattern. In Greece and New Zealand, virtually all educational funding is done by the central level of government, while in Belgium, Japan, Korea, Switzerland, and the United States, less than 1 percent of funding is undertaken by the central government. In Belgium and Korea, virtually all funding is done by the regional level of government. Local governments allocate virtually all final funds in the United States, and the overwhelming part of funds in Canada, Hungary, Norway, and the United Kingdom.

Tables 4 and 5 report the effect of the extent of decision-making at the central level of government ("central government decisions"), where the residual category (the percentage of decisions not taken at the central level) encompasses the decisions taken at the school level and at the local and

---

[27] In Greece, all decisions on personnel management are taken at the central level. In Portugal, all decisions on planning and structures are taken at the central level.

intermediate (sub-regional, provincial, and state) levels of government. Students in countries with a higher percentage of decisions taken at the central level of government scored lower in both the mathematics and the science tests, with only the science effect being statistically significant.

The effects in each of the four sub-groups of decision-making in both mathematics and science are also negative, with the effects of instructional and resource decisions in both mathematics and science and the personnel management effect in science being statistically significant. For example, all other things equal, increasing the amount of decisions on organization of instructions taken at the central level by 1 standard deviation resulted in 0.203 standard deviations lower performance in mathematics (0.134 in science). By contrast, the percentage of decisions taken at an intermediate level of government (part of the residual category in the regressions presented in the tables) were positively related to student performance.

In a similar way to the distribution of decision-making authority, the distribution of responsibility for and control over funding between the different government levels is related to student performance. The larger the share of funds provided at the local or the central level of government in a country (as final purchasers of educational resources),[28] the lower was students' performance in mathematics.[29] Consequently, students performed considerably better the more funding was decided on at an intermediate level of government (the residual category). Once responsibility for decision-making in and funding of education lies with the administration, an administrative level distant enough from individual schools to limit opportunistic and collusive behavior but not as remote as the center, where

[28] Qualitatively similar results arise when looking at the level of government being the initial source of funds, i.e. before transfers between the different government levels.

[29] The effects in science are statistically insignificantly negative for centrally provided funds and virtually zero for locally provided funds.

decision-makers are no longer familiar with local needs, seems to be most conducive to focusing attention on student performance.

Private Schools

The extent of competition from private institutions in the education system differs considerably across countries, leading to differences in the market structure prevailing in the education systems. The Netherlands have by far the highest share of privately managed schools (76 percent of all schools), followed by the United Kingdom (36 percent) and Korea (35 percent). However, less than one percent of Dutch schools are financially independent in the sense that they receive less than half of their core funding from government agencies. The countries with the largest shares of financially independent private schools are Japan (24 percent), Korea (18 percent), and the United States (16 percent). At the other extreme, Australia, Austria, the Czech Republic, Denmark, France, Germany, Hungary, Iceland, Norway, Spain, and Sweden have virtually no independent private schools. As the results in Tables 4 and 5 show, students in countries with larger shares of enrollment in privately managed educational institutions ("private enrollment") scored statistically significantly higher in both mathematics and science. That is, countries with a higher share of private management control over schools performed better. This effect was even larger when only those private institutions were considered which were also financially independent of funding from government sources for their basic educational services ("independent private enrollment").

The Netherlands (75 percent) and Belgium (63 percent) are the countries with by far the largest share of public educational expenditure going to private institutions, while in Austria, Greece, Ireland, New Zealand, the Russian Federation, and the United States, less than half a percent of public expenditure goes to private schools. Similar to the results for private enrollment, countries with a higher share of (public) educational expenditure going to private institutions ("public expenditure on private institutions") performed better both in mathematics and in science (with only the mathematics effect being statistically significant). Again, this effect was even stronger when only those expenditures were counted which went to independent private institutions which received less than half of their core funding from the government ("public expenditure on independent private institutions"). Thus, student performance is higher in education systems where private schools take over resource allocation from public decision-makers.

These effects of private school management are measured at the country level. This does not allow for an assessment of the relative performance of public and private schools, for which the relevant data is not available in the TIMSS case. However, measuring the system-level effect of private school management is the appropriate way to estimate the general effects of the competitive environment and the market structure prevailing in the different education systems, because increased competition from private schools is expected to have a positive effect also on the effectiveness of resource use in nearby public schools which may otherwise lose students to the private schools. As Hoxby (1994) has shown, increased competition from private schools statistically significantly raises the performance of public school students in US metropolitan areas, so that positive effects of private school competition on nearby public schools are clearly given. Furthermore, Hoxby (1996) has shown that the negative effect of teacher unionization is statistically significantly reduced in the United States when a school faces competition from private schools. Hoxby suggests that teacher unions may actually be a primary means whereby a lack of competition among schools translates into more generous school inputs and worse student performance.[30] In any case, a structure of the education system characterized by competition through private institutions appears to be highly conducive to students' learning, as the performance of students is found to be statistically significantly positively related to the share of private management control over educational institutions.

4.4 Summary of Microeconometric Results and Robustness of Data Imputation

Overall, the microeconometric results suggest that both family and institutional factors have strong and unambiguous effects on students' educational performance, while the impact of resource factors appears to be dubious and weak at best. The students' family background exerts the strongest impact on students' performance. Especially the educational background at home has strongly conducive effects on students' learning in mathematics and in science. By contrast, resource effects are very weak, and many even imply that a superior provision with resources can go hand in hand with inferior student performance. There is certainly no general positive effect of educational spending on educational outcome. Smaller student-teacher ratios in the classroom or at the school level do not in general lead to increased student performance. However, a sufficient equipment with instructional material and a sufficient level of teachers' formal education seem to render positive effects.

---

[30] As direct evidence on the relative performance of private schools in the United States, Rouse (1998) has shown that gains in mathematics scores of students who participated in the Milwaukee Parental Choice Program to attend nonsectarian private schools were statistically significantly higher than those of unsuccessful applicants to the program and of other public school students, controlling for student fixed effects. A collection of contributions to the discussion whether private schools are superior to public schools is contained in Cohn (1997, Part II). Concerning competition among public schools, Hoxby (2000) has shown that easier choice among public schools in US metropolitan areas leads to greater productivity of these schools, both in the form of improved student performance and of lower expenditure per student.

Taken together, the features of the institutional system governing the education process strongly influence students' performance. The effects of the dummies characterizing institutional settings in Tables 2 and 3 sum up to more than 210 test score points in mathematics and to about 150 test score points in science. That is, a student who faced institutions that were all conducive to student performance would have scored more than 200 points higher in mathematics than a student who faced institutions that were all detrimental to student performance. In addition to that, there are the effects of the discrete variables and the system-level results reported in Tables 4 and 5.

Table 6 summarizes the individual institutional effects and relates them to the theoretical hypotheses of institutional economics applied to the education sector of Section 2. Central examinations as well as centralized control mechanisms with respect to standards and budgets seem to favor educational performance. School autonomy favors educational performance in the domains of personnel management like choice and rewarding of teachers and of process decisions like purchase of supplies. Given that responsibility for decision-making in and funding of education lies with the administration, intermediate levels seem to fare better than both central and local authorities in focusing attention on student performance. A competitive environment in the education sector characterized by large shares of private school management helps to assure that the producers of education look for the performance of students. An individual teacher having decision-making power over her teaching methods and devoting extra time to student assessment is conducive to students' learning. By contrast, a strong influence of teacher unions - or more generally teachers acting collectively - in the education system seems to have a negative impact on student performance.

Table 6: Theoretical hypotheses and microeconometric results on institutional effects

	Influence on student performance	
	Theoretical hypothesis (Sections 2.2/2.3)	Microeconometric evidence (Section 4.3)
Central examinations	+	+
Central control of standards	+	+
School autonomy in budgetary matters	–	–
School autonomy in personnel management	+	+
School autonomy in process decisions	+	+
Intermediate level of administration and funding	+ / –	+
Private school management	+	+
Individual teachers' influence on teaching methods	+ / –	+
Teacher unions' influence on curriculum	–	–
Scrutiny of student assessment	+	+
Parents' influence	+	+ / –

+ = positive impact. – = negative impact. + / – = ambiguous impact.

A comparison between performance in mathematics and in science shows that all of these results are very robust across the two subjects. Family and resource effects as well as institutional effects are qualitatively the same for mathematics and science learning. The only difference is that standardization effects seem to be more positive in mathematics than in science. This shows up in the facts that the effects of centralized examinations, curricula, and textbook approval are larger for mathematics than for science, that a strong influence of external examinations on the school curriculum has a positive effect on mathematics scores but an ambiguous one on science scores, and that school authority in the four decision-making domains reported by the OECD impacts positively on science performance but is unrelated to mathematics performance (with the exception of the organization of instructions). This difference may indicate a higher propensity for standardization in the case of mathematics than in the case of science.

Since some of the variables of the TIMSS data set included a substantial amount of missing values and therefore had to be imputed, it remains to be tested whether the reported results are sensitive to the imputation. The robustness can be tested by dropping observations with imputed data individually for each variable and re-running the regressions. The only changes either in significance or direction of the relationships occur in the regressions for the following institutional variables.[31] The effect of external exams' influence on the curriculum turns positive (albeit statistically insignificant) in science, replicating the mathematics result. The negative effect of teachers' responsibility for the school budget turns strongly statistically significant in science, while it is statistically significant only at the 15 percent level in mathematics. The coefficient on subject teachers' influence on the curriculum turns statistically insignificant and positive in both mathematics and science, as does the coefficient on school teachers' influence on the curriculum in science. The coefficient on the class teacher having strong influence on the subject matter taught turns statistically significant in science, while the insignificant coefficient on the choice of textbooks turns positive in science (as it is in mathematics). The effect of homework in mathematics and the effect of parents' influence on the

[31] The only changes for background and resource effects are that the coefficient on parents having some secondary education turns statistically insignificant both in mathematics and in science (and positive in mathematics), that the coefficient on the community location close to the center of a town turns statistically insignificantly positive in science and statistically significantly positive in mathematics, and that the positive effect of the teacher's highest education level being the BA turns statistically insignificant in the mathematics regression.

curriculum in both mathematics and science turn statistically insignificantly negative. Since the negative impact of teachers exercising collective influence over the curriculum is anyway best represented by the strong negative effect of teacher unions, and since a statistically significant impact of parents' involvement in teaching is shown by the strong negative impact of uninterested parents, it can in sum be stated that none of the findings relevant for the argumentation in this study depend on the data imputation. Furthermore, increasing the threshold of non-imputed variables for a student to be included in the sample (see Section 3.3) by another 10 variables - reducing the total sample size to 255,018 students in mathematics and to 251,292 students in science - does not lead to any change in significance or sign of the coefficients.

5 Understanding Cross-Country Differences in Student Performance

5.1 Macro Education Production Functions

In their country-level regressions of test scores on various measures of family and school inputs, Hanushek and Kim (1999) and Lee and Barro (1997) achieve very low explanatory power as measured by the adjusted R^2 (the proportion of the variation in test scores explained by explanatory variables). Hanushek and Kim report R^2s between 0.17 and 0.25 for different estimations, while the average R^2 of Lee and Barro's panel estimation is 0.23. In microeconometric (student-level) estimations like the ones reported in Section 4, the omission of measures of innate ability leads to relatively low R^2s since unobserved heterogeneity in the innate ability of students enters the error term ε. However, assuming that the average level of innate ability does not vary across countries (as Hanushek and Kim explicitly do), the proportion of the variation in test scores explained by measurable variables should be considerably higher in country-level regressions than in microeconometric regressions. Put differently, the low R^2s of the available country-level studies reveal that their measures of family and school inputs cannot explain much of the cross-country differences in the performance levels of students.

If the hypotheses presented in Section 2 are correct and institutional features are important in determining student performance, incorporating aggregated institutional variables into a macro education production function should increase the explained proportion of the variation in test scores considerably. By aggregating institutional variables to reflect the percentage of students in a country for whom an institutional feature is given[32] and by combining this institutional data with country-level data on average test scores, family background, and resource endowment for each country, one can devise a rough-and-ready method for estimating institutional effects at the country level. In a country-level estimation, a lot of the valuable information of the student-level estimation, which links the teaching environment of each individual student directly to her performance, is lost. This forgoes much of the valuable insights attainable in the microeconometric analysis. But performing a country-level estimation yields a test whether institutions matter for the observed cross-country differences in the performance of students and provides some indication of the explanatory power of institutional features.

5.2 Country-Level Results

Table 7 reports the results of such a country-level regression for student performance in mathematics and in science. The regressions include the

---

[32] The aggregation of the variables takes care of the stratification of the TIMSS data by weighting each student to yield nationally representative aggregated data for each country.

share of parents who finished secondary education and the share of parents who had education beyond secondary education[33] as indicators of the family background of the students of each country, average class size as a resource indicator, and three indicators of institutional features of the education systems. The first institutional indicator is the extent of school autonomy in supply choice in the country as measured by the percentage of students in schools which had primary responsibility for purchasing supplies. The second one is an indicator of teacher unions' influence as measured by the percentage of students in schools where teacher unions had a strong influence on the curriculum. The third institutional indicator is a proxy for the scrutiny of assessment as measured by the average time class teachers spent outside the formal school day on preparing or grading student tests or exams. The aggregated (country-level) variables give the average manifestation of a characteristic in a country or the percentage of students in the country for which a characteristic is given, while micro (student-level) variables directly link the characteristic to each student. To emphasize this difference between micro and aggregated variables, the aggregated variables of Table 7 have names adapted but different from the micro variables.

The indicators of family background have a strong positive impact on the average educational performance of the students in a country. The share of parents who finished secondary education has the largest standardized coefficient both in the mathematics and in the science regression. It actually appears that a completed basic education of the parents at the secondary level is far more important for students' learning than parents having education beyond the secondary level.

33 To save on degrees of freedom, the latter share combines the information contained in the two dummies in the student-level estimation on parents who had some education beyond the secondary level and on parents who finished university.

Table 7: Country-level results for student performance

Dependent variable: TIMSS international mathematics/science test score.
Standard errors in parentheses.

I. Mathematics	Coeff.	S.E.	Std. Coeff.
Constant	144.424†	(63.234)	
Parents' education: finished secondary	176.271*	(23.308)	0.725
Parents' education: beyond secondary	91.367*	(22.557)	0.429
Class size	3.873*	(0.720)	0.524
School autonomy in supply choice	98.464‡	(49.381)	0.170
Teacher unions' influence	-467.790*	(90.000)	-0.455
Scrutiny of assessment	26.911*	(6.610)	0.355
Observations	39		
F	19.31		
R^2 (adj.)	0.74		

II. Science	Coeff.	S.E.	Std. Coeff.
Constant	234.843*	(62.028)	
Parents' education: finished secondary	136.873*	(23.448)	0.705
Parents' education: beyond secondary	60.989†	(23.174)	0.359
Class size	2.557*	(0.856)	0.370
School autonomy in supply choice	106.307†	(48.440)	0.230
Teacher unions' influence	-369.327*	(92.859)	-0.450
Scrutiny of assessment	12.428‡	(6.963)	0.197
Observations	39		
F	10.49		
R^2 (adj.)	0.60		

* Significant at the 1 percent level.
† Significant at the 5 percent level.
‡ Significant at the 10 percent level.

As an indicator of the amount of resources used in the education system, average class size is again statistically significantly positively related to student performance. The test score levels in mathematics and science were higher in education systems with larger classes, unambiguously implying that resources are more effectively used in countries with larger classes. As an alternative measure of resource effects, expenditure per student entered statistically insignificantly into the equation both in mathematics and in science. Thus, higher resource use in schooling clearly does not contribute to an explanation of international differences in educational performance levels. This finding on resource effects in the TIMSS data set corroborates the evidence presented by Hanushek and Kim (1999), where the estimated resource effects are also either statistically insignificant or statistically significant but with the wrong sign. However, this finding contrasts with the finding of Lee and Barro (1997) that smaller class sizes are related to superior student performance.[34]

The three institutional indicators included in the country-level regressions are all statistically significantly related to student performance in both mathematics and science. School autonomy in supply choice is positively related to student performance, indicating that decentralized decision-making on process decisions is beneficial to teaching outcomes. The indicator of teacher unions' influence in the education system shows a strong negative effect on student performance in both mathematics and science. An increase of 1 standard deviation in the teacher union indicator is related to a decrease of 0.46 standard deviations in the mathematics test score (0.45 in science). The proxy for scrutiny of assessment has a positive impact on the student performance, which is larger in mathematics than in science.

The six variables together explain the international variation in student test scores far better than previous studies. They yield an adjusted R^2 of 0.74 in mathematics and of 0.60 in science. This indicates that even with the rough-and-ready country-level method, where the indicators can only proxy for actual features of the education systems across the 39 countries but cannot link them microeconometrically to individual student performance, three quarters of the cross-country variation in mathematics test scores and

34 As their measure of class size, Lee and Barro (1997) use the average student-teacher ratio in primary schools of a country, while their test scores mainly reflect performance in secondary education. This measurement seems inferior to using the actual class sizes of the students tested, as was done in this study. Lee and Barro also report a statistically insignificant negative coefficient on expenditure per student.

60 percent of the variation in science test scores can be explained when including the three indicators of institutional features of the education systems.

Cross-country differences in student performance are not a mystery. They are related to policy measures. However, the policy measures which matter for schooling output are not simple resource inputs. Instead, differences in the institutions of the education systems help in understanding cross-country differences in students' educational performance. Institutions set incentives and thereby determine the effectiveness with which schooling resources are put to use. Success in educational production does not primarily depend on the amount of resources spent, but on the institutional features governing the education process.

6 Conclusions

Microeconometric student-level estimates reveal that differences in the incentive structures determined by the institutional features of the education systems matter for student performance. The combined effect of performance-conducive educational institutions amounts to a test score difference of more than 200 points in mathematics (150 points in science), which equals about 2 standard deviations in test scores and compares to an average test score difference between seventh and eighth grade of 40 points. Aggregated country-level estimates show that institutional differences can explain a major part of the international differences in average student performance levels. They help to explain 74 percent of the cross-country variation in mathematics test scores and 60 percent in science test scores in this study, whereas previous studies which constrained themselves to family and resource effects reached only a maximum of 25 percent of explained variation.

For education policy, this means that the crucial question is not one of more resources but one of improving the institutional environment of education to ensure an efficient use of resources. Student performance is influenced by the productivity of resource use in schools. This productivity is determined by the behavior of the people who act in the educational process. These people respond to incentives. And their incentives are set by the institutional structure of the system. In short, by setting proper institutions, education policy can favorably affect student performance. By contrast, spending more money within an institutional system that sets adverse incentives will not improve student performance. The only policy that promises positive effects is to create an institutional system where all the people involved have an incentive to improve student performance.

The empirical results identify the specific institutional features of the schooling system which are favorable to student performance. Among these features are:

- central examinations,
- centralized control mechanisms in curricular and budgetary affairs,
- school autonomy in process and personnel decisions,
- an intermediate level of administration performing administrative tasks and educational funding,
- competition from private educational institutions,
- individual teachers having both incentives and powers to select appropriate teaching methods,
- limited influence of teacher unions,
- scrutiny of students' educational performance, and
- encouragement of parents to take interest in teaching matters.

Once these favorable institutional features are implemented in a country, this should allow considerable improvements in the educational performance of students, conceivably at the given size of the educational budget. One of the most pleasant features of institutional reforms - as opposed to resource expansions - is in fact that institutional effects are "for given resources." That is, they mostly come at no or low recurrent operation costs, causing only implementation costs which usually lie in the political domain. Coase (1984, p. 230) once stressed that "[t]he choice in economic policy is a choice of institutions." Education policy is not different.

These results open a wide array of questions to be dealt with in future research. If it is a choice of institutions that matters in education policy, then what determines the outcome of a given institutional choice? What determines the evolution of educational institutions given the multitude of different groups of people with differing interests and of organizations reflecting these interests? An understanding of how institutions can be changed requires the endogenization of the rules and institutions governing the education system. To achieve this, a theoretical model is required which represents the behavior of the different educational agents in response to institutionally set incentives and replicates the empirical evidence on students' performance presented in this paper.

Furthermore, the estimation of a more complex empirical model of educational production could yield additional findings. Such a model could be represented in a system of equations where both educational performance and the amount, composition, and effectiveness of the resources used are a function of particular institutional features of the education system. In the light of the arguments raised in this paper, the observed adverse effects of expenditure and class size on student performance might actually to some extent reflect the power of organized teachers. That is, more resources are spent on education when teachers have collective power, which in turn may be responsible for the poor effectiveness of resource use. While this would give a reason for the systematic relation of improved resource availability to inferior resource effectiveness and thus inferior student performance, more detailed research is clearly warranted on this issue.

Empirical research could also further scrutinize the interaction between different institutional features of the education system by including interaction terms in the estimation. Additionally, the wealth of information given in the TIMSS data base allows for an empirical analysis in other fields than economics. Educationists might be interested in the effect of different instructional modes on students' performance, possibly differentiated by the institutional setting in which they are applied. Sociologists might analyze peer effects on individual students' educational achievement. While this study has focused on test scores as educational outcomes, it should also be of interest whether the institutional features conducive to student performance in mathematics and science have (positive or negative) effects on the social behavior of students such as use of violence or attitudes towards societal values.

Other important empirical topics include the impact of performance-related pay for teachers and of voucher systems on student performance. The negative effect of increased influence of teacher unions found in this paper might actually proxy for the effect of a standard salary scale as opposed to merit differentials in teacher pay. Since neither performance-related pay nor voucher systems have yet been implemented on a wide-ranging scale, their effects cannot be estimated with the TIMSS data. However, case-study evidence may increasingly become available and complement the analysis of this paper.

References

Angrist, Joshua D., Victor Lavy (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics* 114(2): 533-575.

Beaton, Albert E., Michael O. Martin, Ina V.S. Mullis, Eugenio J. Gonzalez, Teresa A. Smith, Dana L. Kelly (1996a). *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS)*. Boston College, Chestnut Hill, MA.

Beaton, Albert E., Ina V.S. Mullis, Michael O. Martin, Eugenio J. Gonzalez, Dana L. Kelly, Teresa A. Smith (1996b). *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS)*. Boston College, Chestnut Hill, MA.

Bishop, John H. (1989). Is the Test Score Decline Responsible for the Productivity Growth Decline? *American Economic Review* 79(1): 178-197.

Bishop, John H. (1992). The Impact of Academic Competencies on Wages, Unemployment, and Job Performance. *Carnegie-Rochester Conference Series on Public Policy* 37: 127-194.

Bishop, John H. (1997). The Effect of National Standards and Curriculum-Based Exams on Achievement. *American Economic Review, Papers and Proceedings* 87(2): 260-264.

Bishop, John H. (1999). Are National Exit Examinations Important for Educational Efficiency? *Swedish Economic Policy Review* 6(2): 349-398.

Boissiere, M., J.B. Knight, R.H. Sabot (1985). Earnings, Schooling, Ability, and Cognitive Skills. *American Economic Review* 73(5): 1016-1030.

Bryk, Anthony S., Valerie E. Lee (1992). Is Politics the Problem and Markets the Answer? An Essay Review of Politics, Markets, and America's Schools. *Economics of Education Review* 11(4): 439-451.

Bryk, Anthony S., Stephen W. Raudenbush (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Advanced Quantitative Techniques in the Social Sciences 1. Sage Publications, Newbury Park et al.

Card, David, Alan B. Krueger (1992). Does School Quality Matter? *Journal of Political Economy* 100: 1-40.

Case, Anne, Angus Deaton (1999). School Inputs and Educational Outcomes in South Africa. *Quarterly Journal of Economics* 114(3): 1047-1084.

Case, Anne and Motohiro Yogo (1999). Does School Quality Matter? Returns to Education and the Characteristics of Schools in South Africa. *NBER Working Paper* 7399, October.

Chubb, John E., Terry M. Moe (1990). *Politics, Markets, and America's Schools*. The Brookings Institution, Washington, D.C.

Coase, Ronald H. (1984). The New Institutional Economics. *Journal of Institutional and Theoretical Economics* 140: 229-231.

Cohen, Jon, Stéphane Baldi (1998). An Evaluation of the Relative Merits of HLM vs. Robust Linear Regression in Estimating Models with Multi-Level Data. Mimeo, American Institutes for Research, February.

Cohn, Elchanan (ed.) (1997). *Market Approaches to Education. Vouchers and School Choice*. Pergamon/Elsevier Science, Oxford.

Costrell, Robert M. (1994). A Simple Model of Educational Standards. *American Economic Review* 84(4): 956-971.

Deaton, Angus (1997). *The Analysis of Household Surveys*. The Johns Hopkins University Press, Baltimore and London.

DuMouchel, William H., Greg J. Duncan (1983). Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples. *Journal of the American Statistical Association* 78(383): 535-543.

Federal Reserve Bank of New York (FRBNY) (1998). Excellence in Education: Views on Improving American Education. *Economic Policy Review* 4(1).

Furubotn, Eirik G., Rudolf Richter (1997). *Institutions and Economic Theory. The Contribution of the New Institutional Economics.* The University of Michigan Press, Ann Arbor.

Goldstein, Harvey (1999). *Multilevel Statistical Models*. Kendall's Library of Statistics 3, Revised internet edition of the 2nd print edition, available at http://www.arnoldpublishers.com/support/goldstein.htm. Edward Arnold, London.

Gonzalez, Eugenio J., Teresa A. Smith (eds.) (1997). *User Guide for the TIMSS International Database* - *Primary and Middle School Years*. International Association for the Evaluation of Educational Achievement, TIMSS International Study Center, Boston College, Chestnut Hill, MA.

Gundlach, Erich, Ludger Wößmann (1999). The Fading Productivity of Schooling in East Asia. *Kiel Working Paper* 945, September.

Gundlach, Erich, Ludger Wößmann, Jens Gmelin (2001). The Decline of Schooling Productivity in OECD Countries. *Economic Journal*: forthcoming.

Hanushek, Eric A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature* 24: 1141-1177.

Hanushek, Eric A. (1995). Interpreting Recent Research on Schooling in Developing Countries. *World Bank Research Observer* 10(2): 227-246.

Hanushek, Eric A. (1996). School Resources and Student Performance. In: Gary Burtless (ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. The Brookings Institution, Washington, D.C.: 43-73.

Hanushek, Eric A. and Associates (1994). *Making Schools Work. Improving Performance and Controlling Costs*. The Brookings Institution, Washington, D.C.

Hanushek, Eric A., Dongwook Kim (1999). Schooling, Labor Force Quality, and Economic Growth. Mimeo, University of Rochester, October. (Revised Version of *NBER Working Paper* 5399).

Heckman, James J. (1999). Policies to Foster Human Capital. *NBER Working Paper* 7288, August.

Heckman, James J., Anne Layne-Farrar, Petra Todd (1996). Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings. *Review of Economics and Statistics* 78(4): 562-610.

Hedges, Larry V., Richard D. Lane, Rob Greenwald (1994). Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes. *Educational Researcher* 23(3): 5-14.

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

Hoxby, Caroline M. (1996). How Teachers' Unions Affect Education Production. *Quarterly Journal of Economics* 111: 671-718.

Hoxby, Caroline M. (1999). The Productivity of Schools and Other Local Public Goods Producers. *Journal of Public Economics* 74: 1-30.

Hoxby, Caroline M. (2000). Does Competition among Public Schools Benefit Students and Taxpayers? *American Economic Review*: forthcoming.

International Association for the Evaluation of Educational Achievement (IEA) (1998). *Third International Mathematics and Science Study. International Achievement Reports.* Available from: http://wwwcsteep.bc.edu/TIMSS1/Achievement.html.

Journal of Economic Perspectives (JEP) (1996). Symposium on Primary and Secondary Education. *Journal of Economic Perspectives* 10(4).

Klein, Lawrence R. (1974). *A Textbook of Econometrics*. Second Edition. Prentice-Hall, Englewood Cliffs, N.J.

Krueger, Alan B. (1999). Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics* 114(2): 497-532.

Landsburg, Steven E. (1993). *The Armchair Economist: Economics and Everyday Life*. The Free Press/Macmillan, New York et al.

Lee, Jong-Wha, Robert J. Barro (1997). Schooling Quality in a Cross Section of Countries. *NBER Working Paper* 6198, September.

Martin, Michael O., Dana L. Kelly (eds.) (1998). *TIMSS Technical Report Volume II: Implementation and Analysis, Primary and Middle School Years*. Boston College, Chestnut Hill, MA.

Murnane, Richard J., John B. Willett, Frank Levy (1995). The Growing Importance of Cognitive Skills in Wage Determination. *Review of Economics and Statistics* 77: 251-266.

North, Douglass C. (1994). Economic Performance Through Time. *American Economic Review* 84(3): 359-368.

Organisation for Economic Co-Operation and Development (OECD) (var. iss.). *Education at a Glance. OECD Indicators*. Paris.

Pritchett, Lant, Deon Filmer (1999). What Education Production Functions Really Show: A Positive Theory of Education Expenditure. *Economics of Education Review* 18: 223-239.

Review of Economics and Statistics (RES) (1996). Symposium on School Quality and Educational Outcomes. *Review of Economics and Statistics* 78(4).

Rouse, Cecilia Elena (1998). Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program. *Quarterly Journal of Economics* 113: 553-602.

Shleifer, Andrei (1998). State versus Private Ownership. *Journal of Economic Perspectives* 12(4): 133-150.

Tyler, John H., Richard J. Murnane, John B. Willett (1999). Do the Cognitive Skills of School Dropouts Matter in the Labor Market? *NBER Working Paper* 7101, April.

UNESCO (var. iss.). *Statistical Yearbook.* Paris.

World Bank (1999). *World Development Indicators CD-Rom.*