

THE EFFECT HETEROGENEITY OF CENTRAL
EXAMS: EVIDENCE FROM TIMSS,
TIMSS-REPEAT AND PISA

LUDGER WOESSMANN

CESIFO WORKING PAPER NO. 1330

CATEGORY 4: LABOUR MARKETS

NOVEMBER 2004

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the CESifo website:* www.CESifo.de

THE EFFECT HETEROGENEITY OF CENTRAL EXAMS: EVIDENCE FROM TIMSS, TIMSS-REPEAT AND PISA

Abstract

This paper uses extensive student-level micro databases of three international student achievement tests to estimate heterogeneity in the effect of external exit exams on student performance along three dimensions. First, quantile regressions show that the effect tends to increase with student ability. But it does not differ substantially for most measured family-background characteristics. Second, central exams have complementary effects to school autonomy. Third, the effect of central exit exams increases during the course of secondary education, and regular standardised examination exerts additional positive effects. Thus, there is substantial heterogeneity in the central exam effect along student, school and time dimensions.

JEL Code: I2.

Keywords: central exams, student achievement, international education production function, effect heterogeneity, TIMSS, PISA.

*Ludger Woessmann
Ifo Institute for Economic Research
at the University of Munich
Poschingerstr. 5
81679 Munich
Germany
woessmann@ifo.de*

Financial support from the Volkswagen Foundation for part of the underlying research is gratefully acknowledged, as are the constructive comments by two anonymous referees.

1. Introduction

More and more evidence is accumulating that the existence of central exit exams is strongly positively related to students' academic performance (cf. Bishop, 2004, for a survey). Cross-country studies have shown that students performed better in countries with curriculum-based external exit exams on the 1991 International Assessment of Educational Progress (IAEP) math, science and geography tests, the 1991 International Association for the Evaluation of Educational Achievement (IEA) Reading Literacy study, the 1995 Third International Mathematics and Science Study (TIMSS), the 1999 TIMSS-Repeat study and the 2000 Programme for International Student Assessment (PISA) reading, math and science tests. Likewise, cross-regional studies in countries where some regions have external exit exams and others not have shown the same result for Canadian provinces, US states and German states.¹ Most of these studies implicitly assume a homogeneous effect of central exams and stress their general impact on the behaviour of students (cf. Bishop, 2004; Hanushek, 2002).

Instead, this paper argues that central exams exert heterogeneous performance effects, depending on students' backgrounds, on schools' specific settings and on time patterns, because they change the behaviour of students, parents, teachers and schools in varying ways. For example, the effect of central exams may differ depending on student ability, parental background, the degree of school autonomy and the grade in which they are implemented.

The paper tests the hypothesis of effect heterogeneity using extensive micro data at the student level from the three largest and most recent international student achievement tests, namely TIMSS (conducted in 1995), TIMSS-Repeat (1999) and PISA (2000). Since in nearly all countries, the existence of central exit exams is a national characteristic, the effect of central exams can be estimated best using such international data. Furthermore, in order to detect effect heterogeneity, micro data containing extensive information on student and school background characteristics is necessary. This makes the international micro databases uniquely capable to provide evidence on the heterogeneity of the effects of central exams.

The paper is structured as follows. Section 2 sketches the theoretical background of why the effect of central exams may be heterogeneous along different dimensions. Section 3 briefly describes the databases, as well as the basic empirical specification. Section 4 presents empirical evidence on effect heterogeneity of central exams along the student dimension,

¹ The specific references for the different studies are: IAEP: Bishop (1995, 1997); IEA Reading: Bishop (1999); TIMSS: Bishop (1997), Wößmann (2003a); TIMSS-Repeat: Wößmann (2003b); PISA: Bishop (2004), Fuchs and Wößmann (2004); Canadian provinces: Bishop (1995, 1997, 1999); US states: Bishop (1995), Bishop et al. (2000, 2001); German Länder: Jürges et al. (2003).

including both student ability and family background. Section 5 tests effect heterogeneity along the school dimension, focussing on differential effects by school autonomy. Section 6 presents evidence on effect heterogeneity along the time dimension, considering both whether the effect of central exit exams differs through the course of secondary education and whether more regular standardised testing throughout secondary education renders additional effects. Section 7 concludes with a summary of the main findings.

2. Theoretical Background

It is the impact of information on incentives and behaviour that forms the theoretical background of this paper. Simple extensions of the basic model of central-exam effects (e.g., Bishop and Wößmann, 2004), adding additional plausible assumptions and additional actors such as parents, teachers and schools, give rise to predictions that the effect of central exams on student performance may be heterogeneous along the student, school and time dimensions.

2.1 Heterogeneity by Student Characteristics

2.1.1 Labour-Market Signalling and Heterogeneity by Ability

There may be several reasons why central exams may affect students of different ability levels differently. For example, central exams may exert non-linear effects because of curricular thresholds, as in minimum competency exams or in exams with particularly high curricular standards. In such curriculum-based arguments, the existence and direction of effect heterogeneity depend on the particular curricular design.

This paper focuses on incentive-based models that consider the signalling value of different kinds of exams on the labour market. Under plausible assumptions, such models predict that central exams may have stronger effects on high-ability students than on low-ability students. Consider the classical model of why central exams may affect student learning (e.g., Bishop, 1995, 2004; Bishop and Wößmann, 2004). This model suggests that grades on centralised exams provide a better signal than locally-graded exams of what students have learnt to employers and to institutions of higher education. Thus, students' rewards for learning are higher where central exams are in place. Due to these increased incentives to learn, student performance will be higher in central-exam systems than in systems without central exams.

Now, let us supplement this basic model by the following two assumptions. First, local employers can assess the signalling value of local grades better than employers from other

regions. That is, locally-graded exams to some extent carry a local signal, but not an inter-regionally assessable signal. Second, low-ability students are less mobile than high-ability students. Low-ability students tend to look for jobs on the local labour market, while high-ability students have a stronger propensity to migrate to other areas to find better job matches.

Given the combination of these two additional assumptions, the model will predict an effect heterogeneity of central exams by student ability. For low-ability students, the positive incentive effect of central exams is relatively small, because locally-graded exams also carry some signal for them. Thus, the additional positive incentive effect of central exams as an inter-regionally assessable signal will be relatively small for them. For high-ability students, by contrast, the locally-graded exams do not have a signal in other regions where many of them will migrate to. Therefore, the incentive effect of central exams is relatively stronger for high-ability students than for low-ability students.

Thus, the model predicts that the effect of central exams increases with students' ability. However, given that there may be other reasons for an effect heterogeneity by ability, e.g., depending on the specific curricular implementation of the central exams, the issue ultimately remains an empirical question, which will be addressed in Section 4.

2.1.2 Incentives and Parental Behaviour

The incentives created by the signal of central exams will not only affect the behaviour of students and employers, but also of parents. As the parental behaviour resulting from the availability of the signal may be expected to differ for parents of different background and interest, the model can straightforwardly be extended to effect heterogeneity of central exams by family background and parental interest.

Where central exams are in place, parents have information on the performance of their children against an established standard and relative to other students in the education system. Therefore, given central exams, parents can not only assess their children's performance against an absolute standard, but they also have some knowledge to decide on who might be responsible for this performance. For example, parents will generally know the performance of some other students in their children's class and the average performance in the country. Thus, in contrast to a system of teacher grading, parents now know whether it is mainly their own child who is doing badly or whether it is the whole class which is performing badly. That is, with central exams, they are in a better position to monitor the performance of students, teachers and schools. Consequently, parents are able to put pressure on students and/or

teachers – whomever they deem responsible for the poor performance of their children.² When teachers grade their students themselves and students get marks relative to their class mean only, parents are not able to observe the performance of the class relative to the country mean and thus have no information on which to base a potential intervention. The existence of the information disseminated to parents by central exams is thus likely to affect the behaviour of both students and teachers in a way that furthers students' academic performance.

In the same way, parents can now monitor the performance of the whole school relative to other schools, and of the administrative entity relative to others. Given that central exams thus impact student performance through the channel of altered parental behaviour, the impact of central exams on student achievement may differ both for different educational and social backgrounds of the students' parents and – related to but apart from that – for parents showing different degrees of interest in their children's achievement.

First, similar to the curriculum-based arguments raised above, it is not clear *ex ante* whether students from different family backgrounds would be affected differently by central exams. On the one hand, better-educated parents may be better equipped and motivated to use the additional information provided by central exams, so that the impact of central exams may be bigger for students from better family backgrounds. On the other hand, central exams may focus teachers' and schools' attention particularly on the performance of students from poorer family backgrounds who do not face strong learning incentives at home. Given central exams, teachers and schools can be additionally rewarded for achieving better performance among these students. Both kinds of effects could also net out, so that there would not be effect heterogeneity of central exams with respect to observed family background.

Second, while central exams provide information to all parents, not necessarily all parents will be willing and able to make use of it. Thus, the impact of central exams might differ depending on how strongly parents care for their children's progress. In central-exam systems, parents who show interest in how much their children learn have a meaningful foundation to intervene and will probably use this opportunity to pressure students and teachers to increase their effort. But parents who are less concerned with their children's

² In a pure sense, this statement is correct only for central exams before the end of schooling, i.e. for non-exit exams, whereas central exit exams provide information only after individual children have left school. However, parents can use the information provided by central exit exams for other students in upper grades in their children's school to get informed about the relative performance of their school and of specific teachers, which will in turn allow them to pressure for better performance in their own children's teaching.

educational performance may not make use of the additional information. Involvement of interested parents in the teaching process may thus be more beneficial in a central-exam system, while this channel may not work with parents who are less interested. Thus, the effect of central exams may be expected to be stronger where parents are strongly interested in their children's progress.

2.2 *Heterogeneity by School Characteristics*

2.2.1 A Principal-Agent Approach to Educational Production

The complexity of the model of incentive effects of information can be increased further by introducing schools into the model. Rather than being mere technical executors, actors in schools exhibit behaviour based on their own interests, which gives rise to agency problems in educational production. Once we allow schools to differ in their characteristics, particularly with respect to the degree of autonomy of schools and teachers, the model predicts interaction effects between central exams and school autonomy, so that the two work as complements.³

The provision of education can be understood as a network of principal-agent relationships in which a principal (e.g. the parents) commissions an agent (e.g. a school director) to perform a service (the education of the children) on behalf of the principal. If such a principal-agent relationship features both asymmetric decentralised information and a constellation of opposing interests between principal and agent, the agency may lead to inefficient provision (cf. Laffont and Martimort, 2002). For if the agent's interests diverge from those of the principal, and if the principal is not fully informed about the agent's real performance, then the agent may pursue his own interests instead of those of the principal, without the latter becoming aware of this behaviour and thus without him being able to sanction it.

Central exams can contribute to mitigating the problem of incomplete monitoring of the actions of the agents in the education system by supplying information about the performance of individual students relative to the national (or regional) student population. They harmonise the incentives of the agents more strongly with the interests of the principal and thus with the objectives of the education system (cf. Wößmann, 2002). They make the performance status of the students visible and comparable for parents, teachers, potential employers and advanced educational institutions, so that better performance can be rewarded (cf. Bishop and Wößmann, 2004). They also prevent that entire areas of knowledge can be

³ This section builds on Wößmann (2003c).

omitted in individual classes without any consequences for marking, and they reveal to parents and school directors whether the teachers are effective in passing knowledge on to their students.

2.2.2 School Autonomy With and Without Monitoring

School autonomy or the decentralisation of decision-making power in various decision-making areas can be understood as such delegation of a task by a principal to agents. In decision-making areas where both divergent interests and asymmetric information are given, incentives and possibilities exist for the agents to act in an opportunistic way without incurring the risk that such behaviour will be noticed and sanctioned.

The danger of opportunism by decentralised decision-makers is thus limited to those decision-making areas in which their interests diverge from the objective of enhancing students' knowledge. This is, for instance, imaginable whenever the decision concerns the financial position or the workload to be fulfilled by the schools. In such cases, it is rational for the school decision-makers to favour their own interests over the promotion of student performance as long as monitoring agencies such as school governors or parents have no information about the actual behaviour of the schools. In view of the decentralised character of educational provision, there is almost always a high degree of information asymmetry about school behaviour. Nevertheless, it can be at least partially overcome by central exams, which supply information about actual performance.

An additional crucial point to be noted is that in many decision-making areas, local decision-makers know much better than a central agency ever could how education services can be most efficiently provided. Thus teachers usually have a local knowledge lead as regards the best way of teaching their specific students a specific subject. This local knowledge lead can make provision by a local agent more efficient than by a central planning authority. But the decisive factor is whether these decision-makers also have the incentive to exploit their local knowledge lead in providing educational services. This will be the case only when others become aware of whether they have made the effort to utilise their local knowledge – i.e., only when information asymmetries are bridged, for instance by central exams.

Figure 1 represents the ensuing effects on performance by school autonomy for various decision-making areas which may be characterised by the presence or absence of incentives for opportunistic behaviour and of local knowledge leads. In those areas where no incentives for opportunistic behaviour exist because the interests of agent and principal do not diverge,

the effects of school autonomy on performance can be easily determined. If local decision-makers have a knowledge lead in such areas, school autonomy has a positive effect on educational performance. This is because the advantages of local decision-making (local knowledge lead) exist, while the disadvantages (opportunistic behaviour) do not. If local decision-makers have no knowledge lead in these areas, there will be no difference between decentralised and centralised decision-making. In both cases, the effects are not heterogeneous with respect to central exams, because by definition there is no risk of opportunistic behaviour which would have to be averted.

< Figure 1 about here >

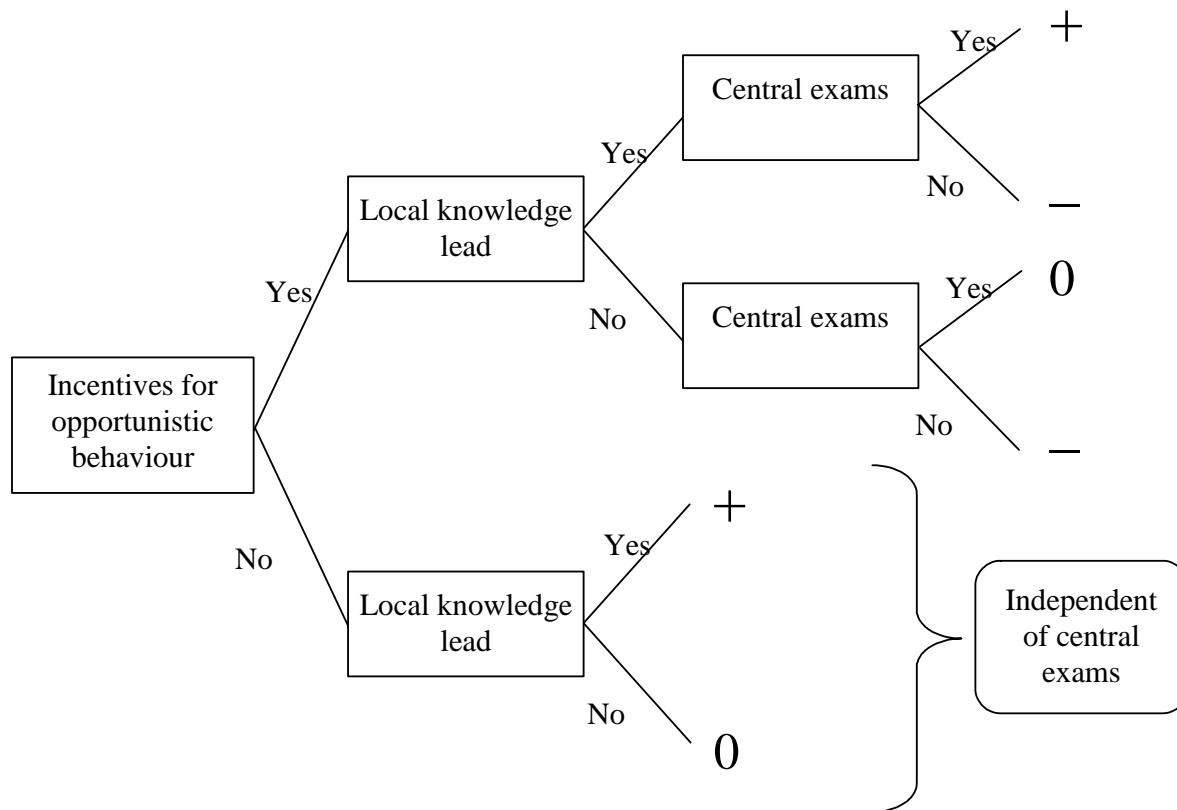
Central exams change the effect of school autonomy on performance only in decision-making areas that offer incentives for opportunistic behaviour due to the diverging interests of agent and principal. In areas without a local knowledge lead and consequently with no benefits of decentralised decision-making, school autonomy has a negative impact on student performance without central exams due to local opportunistic behaviour. But with central exams, the risks of negative performance effects due to local opportunistic behaviour are averted, so that performance will not differ between autonomous and central decision-making.

In decisions containing both incentives for opportunistic behaviour and benefits of superior local knowledge, central exams can avert the disadvantages of opportunistic behaviour, so that the local knowledge lead produces an overall positive effect of school autonomy on performance. Without central exams, the advantage of superior local knowledge must be weighed against the disadvantage of opportunistic behaviour, and the net effect of school autonomy depends on the relative size of these two partial effects. So it is not obvious whether these decision-making areas yield a slightly positive effect, no effect or an overall negative effect of school autonomy. The negative net effect depicted in Figure 1 is based on the empirical results reported in Section 5 below, where the negative opportunism effect tends to outweigh the positive knowledge effect. In this case, central exams turn an originally negative effect of school autonomy on performance completely around to become a positive - effect.

2.3 Heterogeneity by Time

Two simple ideas that extend the model even further give rise to a final dimension of effect heterogeneity, namely the time dimension. The first time aspect considered is whether the effect of central school-leaving exams differs with the grade that students are currently

Figure 1: The Complementarity of Central Exams and Local Autonomy



Notes: “Incentives for opportunistic behaviour” and “local knowledge lead” are features of the respective decision-making area which can be organised either autonomously or non-autonomously.

+: Autonomy enhances student performance. -: Autonomy reduces student performance. 0: No performance difference between autonomous and central decision-making.

Source: Wößmann (2003c).

attending. The second aspect is whether regular standardised testing during secondary education exerts additional effects on top of the effects of central school-leaving exams.

According to the classical model of central-exam effects (cf. Section 2.1.1), central exit exams increase students' incentives to learn because they increase the future external rewards for learning on the labour market. As potential employers and institutions of higher education have central exit exams at their disposal to assess applicants' educational performance, they can base their hiring decisions more on observed educational performance. Therefore, the impact that central exit exams at the end of secondary school exert on student learning may be expected to grow over the course of secondary schooling. The closer students are to the finishing grade, the more important the central-exam effect may get. Thus, there may be a first effect heterogeneity of central exams along the time dimension, in that the effect of central school-leaving exams increases with grade levels.

A second way in which external exams may show heterogeneity along the time dimension is that not only central exit exams, but also more regular central exams during the prior course of schooling may yield positive incentive effects. Regular standardised exams in various grades improve the information status in the education system even in a system that already has central exit exams. Thanks to their earlier availability, in the case of unsatisfactory performance they also allow countermeasures to be taken far ahead of the end of secondary education. Therefore, one might expect regular standardised testing in earlier grades to have positive effects in addition to central exit exams.

3. Data and Empirical Model

3.1 The International Databases

In order to test the different hypotheses derived in the previous section empirically, I use data from three international comparative tests of student performance. The Third International Mathematics and Science Study (TIMSS) was initially carried out in 1995 ("TIMSS-95") and was repeated in 1999 ("TIMSS-Repeat"). TIMSS-95 has internationally comparable data for 266,545 students from 6,107 schools in 39 countries, and TIMSS-Repeat covers 180,544 students from 6,068 schools in 38 countries. The pooled TIMSS database thus contains a total of 447,089 student and 77 country observations. As only 23 countries took part in both tests, the pooled database contains 54 different countries. Both TIMS studies were carried out by the International Association for the Evaluation of Educational Achievement (IEA). In the middle-school years, TIMSS-95 tested the two grades with the largest proportion of thirteen-

year-old students, which in most countries correspond to seventh and eighth grades. TIMSS-Repeat tested only the upper of these two grades. Beaton et al. (1996), Gonzalez and Smith (1997) and Martin and Kelly (1996, 1997) provide detailed information on the TIMSS-95 database (see also <http://timss.bc.edu>). For more information on the TIMSS-Repeat database, see Mullis et al. (2000), Martin et al. (2000) and Gonzalez and Miles (2001).

The third international student achievement test used in this study is the Programme for International Student Assessment (PISA), carried out by the Organisation for Economic Co-operation and Development (OECD) for fifteen-year-old students in 2000. This dataset covers 175,227 students in reading (96,855 in math, 96,758 in science) in 32 countries. Adams and Wu (2002) and OECD (2000, 2001, 2002) provide detailed information on the PISA study (see also <http://www.pisa.oecd.org>). All three studies sampled a representative random sample of schools in each country.⁴

This article uses the individual student data of the three databases, so that as many different education systems with and without central exams as possible can be considered, as well as local differences in student and school characteristics within these systems. In addition to the performance data on math, science and reading of the individual students, the databases contain extensive background information obtained via various questionnaires. Thus, data from student questionnaires allow to control extensively for the influences of the personal and family background of the students in both TIMSS and PISA. In TIMSS, teacher questionnaires contain data on both teacher characteristics and class resources as well as on the influence of teachers in various decision-making areas. Finally, questionnaires of school directors in particular provide information about the degree of school autonomy in various decision-making areas of the TIMSS and PISA schools.

In addition to the data from the student achievement tests, the databases used here contain information about whether central exit exams are held at the end of secondary school in the countries concerned (or in regions within these countries). All forms of “curriculum-based external exit exam systems” (Bishop, 1997) are considered, but not university entrance exams which are not taken by all students and thus do not represent an integral part of the education system. The information about central exams is taken from comparative educational studies, educational encyclopaedias, interviews with representatives of the various national education systems, government documents and background documentation. In cases where central

⁴ Wößmann (2003a, 2002) and Fuchs and Wößmann (2004) contain more detailed information and notes on the specific databases used in this paper.

exams are taken in only some regions of a country, the data used specifies the proportion of students who take them. Unfortunately, internationally comparable data on central exit exams in reading is not available, so that the measure used in this paper is a simple mean of the existence of central exit exams in math and science. Therefore, the results in reading may strongly suffer from attenuation bias due to measurement error.

3.2 *The Basic Empirical Specification*

To estimate the effect α of central exams E on student performance T , the student-level estimations have the following form of an international education production function:

$$(1) \quad T_{ilsc} = \alpha E_c + B_{ilsc} \beta + R_{lsc} \gamma + I_{lsc} \delta + a + \mu_c + \eta_s + \nu_l + \varepsilon_i \quad ,$$

where T_{ilsc} is the test score of student i in class l in school s in country c . E_c is the proportion of students in country c who take part in central exams.⁵ The estimation controls for a host of other variables, including a constant a , vectors of variables for the student's personal and family background B , variables for school resources and teacher characteristics R and variables for several institutional features I such as the extent of school autonomy in various decision-making areas and the centralisation of curricula and textbooks.⁶

The main identifying assumption of this empirical specification is that conditional on all the family, school and institutional background variables, there are no other cross-country features that are consistently related to both central exams and student performance. Then, the least-squares coefficient on central exams will reflect the effect of central exams on student performance. Previous research suggests that the central-exam effect estimated using such an international least-squares specification does not seem to be significantly biased by endogeneity or omitted-variable biases, with neither other institutional differences, nor the general degree of a country's centralisation, nor the homogeneity of a country's population, nor cultural differences driving the estimated effects (Wößmann, 2003b). By using data from three international student achievement tests with different samples of participating countries, this paper can also test for the sensitivity of results to the countries included in each test.

Owing to the complex data structure produced by the survey design of the international tests and the multi-level nature of the explanatory variables, the error term of the regression

⁵ As these involve national central-exam systems in most cases, E is usually assigned the dummy values 0 or 1.

⁶ The specific control variables included for each test differ slightly because of differing availability of background data. For an enumeration of the individual control variables contained in the TIMSS regressions see Wößmann (2002), and in the PISA regressions Fuchs and Wößmann (2004).

has a non-trivial structure. It features components at various levels: μ is a country-specific, η a school-specific, ν a class-specific and ε a student-specific component. The error components are implemented by clustering-robust linear regression (CRLR). The CRLR method relaxes the classical assumption of independence across individual observations and requires only that the observations be independent across primary sampling units (PSUs). This is implemented by using a block-diagonal disturbance variance-covariance matrix with one block for each PSU consisting of the actual disturbance covariance structure of the least-squares disturbances in the PSU (for details, cf. Moulton, 1986; Deaton, 1997, pp. 74-78). By allowing any degree of dependence within PSUs, CRLR lets the data determine the structure of error components in equation (1) to estimate appropriate standard errors in multi-level analyses. That is, in estimating the effect of family and school characteristics, CRLR considers any existing interdependence of the error terms for students within individual schools (and within whole countries when estimating the effect of central exams, which do not vary within countries).

Finally, both TIMSS and PISA used a stratified sampling design within each country, producing varying sampling probabilities. To obtain nationally representative estimates from the stratified survey data, the regressions use weighted least squares (WLS) estimations that weight each student by her sampling probability within countries. This ensures that the proportional contribution to the parameter estimates of each stratum is the same as would have been obtained in a complete census enumeration (DuMouchel and Duncan, 1983; Wooldridge, 2001). Between countries, the weights give equal weight to each country.

4. Evidence on Heterogeneity by Student Characteristics

The following three sections provide empirical evidence on each of the three dimensions of effect heterogeneity of central exams discussed in Section 2. This section analyses how the effect of central exams differs for students with different characteristics (cf. Section 2.1). First, it estimates the effect heterogeneity of central exams on student performance by students' underlying ability level. Second, it estimates interaction effects between central exams on the one hand and family background and parental interest on the other hand.

4.1 Heterogeneity by Student Ability

4.1.1 Quantile Regressions

In order to test whether the effect of central exams on student performance differs by students' ability levels, I estimate equation (1) using quantile regressions (cf. Koenker and Bassett, 1978). The basic learning ability of students – e.g., their innate ability or their learning motivation – remains unmeasured, virtually by definition. However, once all the family-background, resource and institutional effects are controlled for, the conditional performance distribution should be strongly correlated with ability (or, more precisely, with that part of ability that is not correlated with the control variables).

In the following empirical application, this conditional performance distribution is termed “ability.” Quantile regressions estimate the effect of central exams on student performance for students at different points on this ability distribution.⁷

4.1.2 Results on the Effect Heterogeneity of Central Exams by Student Ability

The results concerning the effects of central exams on student performance estimated both by standard least-squares estimations and by quantile regressions are depicted in Figure 2 for the different international datasets. The horizontal lines of Figure 2 show the least-squares coefficients on central exams, estimated using equation (1). On each test, students in countries that have external exit exam systems perform better than students in countries that do not have external exit exam systems. In each case, this effect is statistically significantly different from zero at conventional significance levels.⁸

< Figure 2 about here >

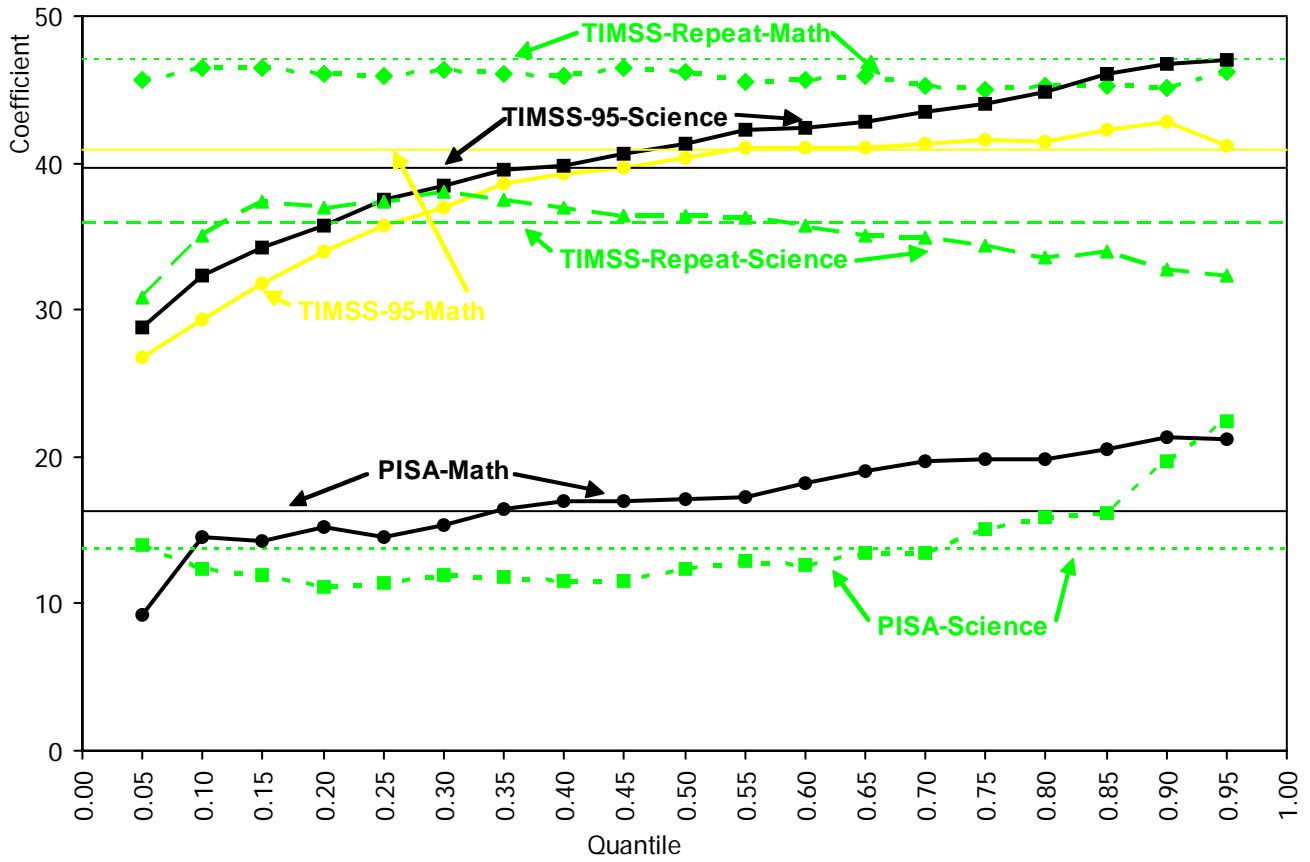
The size of the estimated effects is very large. Both the TIMSS and the PISA test scores are scaled so that each discipline has an international mean of 500 and an international standard deviation of 100 among the participating countries in the respective test. Since the participating countries are different in the different studies, the quantitative results are not directly comparable. In TIMSS-95, the mean difference between seventh and eighth grade in the international sample is roughly 40 test-score points. Thus, the difference in the performance of students in countries with and without central exams is roughly as large as

⁷ Previous education production function studies using quantile regressions are Eide and Showalter (1998) and Levin (2001) for resource effects and Fertig (2003) and Wößmann (2004) for family-background effects.

⁸ While confidence intervals are not shown in the figure for expositional reasons, the standard errors, estimated by clustering-robust techniques using countries as level of clustering, are as follows: TIMSS-95: math 13.5, science 9.9; TIMSS-Repeat: math 13.5, science 12.9; PISA: math 10.2, science 8.2.

Figure 2: Effect Heterogeneity by Student Ability: Quantile Regressions

Coefficient on external exit exams. – Horizontal lines: WLS estimates. –
 Curved lines with dots: Quantile regression estimates for 5% intervals.



Notes: Control variables in the regressions include [number of variables]: TIMSS-95: student and family background [17], resources [14], institutions [18]. TIMSS-Repeat: student and family background [15], resources [14], institutions [18]. PISA: student and family background [40], resources [8], institutions [13].

one grade-level equivalent in the TIMSS studies, i.e. the knowledge learnt in an entire school year. The smaller numerical values estimated for the PISA study are nearly exclusively due to the different scaling according to the different participants in PISA. Once using a rescaled scale that should be more comparable by looking only at countries participating in both studies, the estimated effects in PISA are nearly as large as in the TIMSS studies.

The curved lines with dots in Figure 2 report quantile-regression coefficients of the same model for 19 quantiles ranging from 0.05 to 0.95.⁹ For three of the six considered tests, the effect pattern of central exams on student performance increases with ability: TIMSS-95 math, TIMSS-95 science and PISA math. On these tests, there is effect heterogeneity of central exams by student ability, as predicted in Section 2.1.1. On the PISA science test, the pattern is relatively flat in the lower half (initially slightly decreasing), but it also shows an increasing pattern in the upper half. On the TIMSS-Repeat math test, the pattern is flat, while on the TIMSS-Repeat science test, the pattern is slightly shaped like an inverted U.¹⁰

As indicated before, no decent central-exam information is available in reading, so that the effect is not depicted in Figure 2. However, Fuchs and Wößmann (2004) use a simple mean of the central-exam variable in math and science for the reading regressions and find a positive but relatively small and statistically insignificant effect of 6.9 test-score points. Using the same variable to estimate quantile regressions, I find a pattern of central-exam effects steadily increasing with ability also for PISA reading.

4.2 Heterogeneity by Family Background and Parental Interest

4.2.1 The Interaction Specification

While the above quantile regressions allowed to detect effect heterogeneity along an unobserved dimension, the interest is now in effect heterogeneity along the observed dimension of measures of family background and parental interest. Such kind of heterogeneity can be estimated by adding interaction effects between central exams and the observed family-background characteristics, including measures of parental interest, to equation (1) above:

$$(2) \quad T_{ilsc} = \alpha E_c + (E_c B_{ilsc}) \lambda + B_{ilsc} \beta + R_{lsc} \gamma + I_{lsc} \delta + a + \mu_c + \eta_s + \nu_l + \varepsilon_i \quad .$$

⁹ The statistical precision of the quantile-regression coefficients is slightly lower than that of the least-squares estimates.

¹⁰ One reason for the different pattern in TIMSS-Repeat might be the different set of participating countries.

The estimated interaction effects show whether the effect of central exams on student performance differs for students from different family backgrounds.¹¹

4.2.2 Family Background

Table 1 reports the estimated effects of family background and its interaction with central exams in TIMSS, where the TIMSS-95 and TIMSS-Repeat studies are pooled together. Table 2 reports equivalent results in PISA. The columns labelled “Coefficient” report the coefficient estimate β on the family-background variable itself, and the columns labelled “Interaction” report the coefficient estimate λ on the interaction term between the family-background variable and central exams of the same regression.

< Tables 1 and 2 about here >

Family background generally exerts strong effects on students’ educational performance both in TIMSS and in PISA. In both studies and in all subjects, family-background characteristics such as immigration status, family status, parental education and the number of books at home (serving as a proxy for the educational, social and economic background of the family) are strongly positively related to students’ educational achievement. In PISA, there is additional information on parents’ occupation and work status, which are also significantly related to student achievement.

More importantly for the topic of this paper, there seems to be some, but not much, difference in the family-background effects between systems with and without central exams. In TIMSS, the first heterogeneity is that central exams dampen the effect of the country of birth of students and their parents. That is, immigrants seem to gain more from central-exam systems than nationally born students. Second, the performance advantage of students living with both parents is larger in central-exam systems. Third, central exams also decrease the effect of parental education. Under a system of central exams, it seems to matter less from which parental background students come. Fourth, there is not much evidence that central exams affect students from homes with different amounts of books differently in math. But in science, the positive effect of having more books at home is larger in central-exam systems. This finding counters the effect for parental education, which goes in the opposite direction.

In PISA, there is much less evidence of effect heterogeneity of central exams by family background. First, not a single one of the “books at home” dummies shows a statistically

¹¹ In addition to the family background, resource and institutional controls, the PISA regressions also control for a set of imputation controls (cf. Fuchs and Wößmann, 2004).

Table 1: Interaction Effects between Central Exams and Family Background in TIMSS

	Math		Science	
	Coefficient	Interaction	Coefficient	Interaction
Family background				
Student born in country	20.3 * (2.1)	-4.2 (2.9)	28.1 * (2.1)	-11.5 * (3.0)
Parent born in country	14.8 * (2.1)	-21.6 * (2.8)	11.4 * (2.3)	-8.6 * (2.9)
Living with both parents	5.2 * (1.1)	7.4 * (1.4)	2.6 * (1.0)	7.9 * (1.4)
Parents' education				
Finished primary	22.6 * (1.7)	-14.2 * (2.1)	12.3 * (1.5)	-7.3 * (1.9)
Secondary	31.9 * (1.9)	-17.9 * (2.3)	19.3 * (1.5)	-4.7 + (2.0)
Finished university	49.0 * (2.0)	-15.8 * (2.6)	38.5 * (1.8)	-5.5 + (2.3)
Books at home				
11-25	15.5 * (2.1)	-0.9 (2.4)	15.2 * (1.6)	1.9 (2.0)
26-100	35.6 * (2.3)	6.8 * (2.6)	36.1 * (1.7)	8.0 * (2.2)
101-200	49.9 * (2.4)	4.2 (2.7)	49.0 * (1.8)	11.8 * (2.3)
More than 200	56.7 * (2.5)	4.4 (2.8)	55.9 * (1.9)	13.5 * (2.4)
Parental Interest				
Uninterested parents limit teaching	-9.9 * (3.3)	-7.7 ° (4.1)	-3.1 (2.8)	-16.0 * (3.9)
Interested parents limit teaching	-17.7 * (5.1)	17.8 * (5.9)	-9.8 + (4.1)	17.0 * (5.2)
Grade Level				
Upper grade	25.5 * (0.2)	17.5 * (0.2)	37.0 * (0.1)	12.2 * (0.2)
Students (observations)	447,089		447,089	
Schools (PSUs)	12,175		12,175	
Countries	77		77	
R ²	0.296		0.266	

Dependent variable: TIMSS international math/science test score (pooled TIMSS-95 and TIMSS-Repeat). – Coefficient: Coefficient estimate on the dummy (representing the effect in school systems without external exit exams). – Interaction: Coefficient estimate on the interaction term between the dummy and external exit exams (representing the difference in the effect between school systems without and with external exit exams). – Clustering-robust standard error in parentheses, using schools as level of clustering. – WLS regressions using students' sampling probabilities as weights. – Controls [number of variables]: Central exams [1], background [8], resources [13], institutions [14], institutional interactions [14], centralisation [2]. – Significance levels (based on clustering-robust standard errors): * 1 percent. – + 5 percent.

Source: Based on Wößmann (2002).

Table 2: Interaction Effects between Central Exams and Family Background in PISA

	Math		Science		Reading	
	Coefficient	Interaction	Coefficient	Interaction	Coefficient	Interaction
<i>Born in country</i>						
Student	2.6 (3.4)	2.5 (4.4)	7.9* (2.8)	-1.8 (4.6)	11.4* (2.5)	-5.3 (3.8)
Mother	5.6+ (2.6)	-1.0 (3.5)	11.4* (2.3)	-7.9+ (3.6)	9.3* (2.0)	-3.7 (3.0)
Father	7.9* (2.7)	-5.8 (3.6)	8.9* (2.2)	3.5 (3.5)	5.9* (2.0)	1.3 (2.9)
<i>Living with</i>						
Single father	15.9* (5.3)	1.8 (7.3)	11.4+ (5.0)	15.3+ (7.3)	14.2* (3.9)	2.6 (5.9)
Single mother	6.7 (6.5)	-1.5 (8.8)	8.0 (6.4)	9.7 (9.1)	5.6 (4.9)	2.7 (7.1)
Both parents	13.0+ (5.4)	-0.6 (7.5)	11.1+ (5.1)	12.5° (7.5)	13.9* (4.0)	-0.1 (6.0)
<i>Parents' education</i>						
Primary	5.0 (4.5)	17.5+ (8.5)	10.2+ (4.7)	6.4 (9.8)	13.1* (3.3)	20.4* (6.8)
Lower secondary	8.9° (4.8)	15.2° (8.6)	11.4+ (4.6)	6.4 (9.5)	16.5* (3.4)	14.4+ (7.0)
Upper secondary 1	21.1* (5.0)	1.3 (8.9)	19.9* (4.9)	2.0 (9.9)	27.3* (3.8)	6.6 (7.2)
Upper secondary 2	17.3* (4.7)	13.9 (8.7)	17.2* (4.8)	12.4 (9.8)	25.7* (3.7)	15.6+ (7.0)
University	21.6* (4.8)	14.3 (8.8)	21.8* (4.8)	15.9 (9.8)	28.2* (3.6)	19.2* (7.1)
<i>Parents' work status</i>						
At least one half-time	-3.7 (3.1)	5.3 (4.3)	-3.8 (3.0)	1.4 (4.3)	-6.7* (2.2)	6.9+ (3.2)
At least one full-time	14.8* (2.5)	-1.0 (3.3)	10.7* (2.4)	-1.1 (3.3)	8.2* (1.8)	2.8 (2.5)
Both full-time	11.8* (2.6)	3.2 (3.4)	8.9* (2.6)	3.0 (3.5)	6.3* (1.9)	6.5+ (2.7)
<i>Parents' job</i>						
White collar	15.5* (1.5)	0.7 (1.9)	14.5* (1.2)	1.3 (1.8)	17.1* (1.1)	3.4+ (1.5)
Blue collar	-7.5* (2.0)	-7.3* (2.8)	-10.6* (1.9)	-2.4 (2.8)	-10.1* (1.4)	-6.6* (2.1)
<i>Books at home</i>						
1-10 books	12.8* (4.2)	1.1 (6.2)	16.4* (4.0)	-6.8 (6.4)	21.1* (3.1)	-2.0 (4.9)
11-50 books	25.3* (4.3)	5.0 (6.2)	29.7* (3.8)	-1.5 (6.0)	38.6* (3.2)	0.7 (5.0)
51-100 books	36.2* (4.4)	0.5 (6.4)	39.3* (3.9)	-3.5 (6.1)	48.2* (3.2)	-0.8 (5.0)
101-250 books	52.6* (4.6)	-1.0 (6.5)	51.9* (4.0)	-1.5 (6.2)	62.8* (3.3)	2.0 (5.1)
251-500 books	65.3* (4.7)	-4.9 (6.6)	65.2* (4.1)	-5.6 (6.3)	76.4* (3.4)	-2.2 (5.2)
More than 500 books	68.0* (4.8)	-5.1 (6.8)	64.2* (4.2)	-1.0 (6.4)	75.5* (3.6)	0.5 (5.4)
Students (observations)	96,855		96,758		174,227	
Schools (PSUs)	6,611		6,613		6,626	
Countries	31		31		31	
R ²	0.341		0.288		0.349	

Dependent variable: PISA international math/science/reading test score. – Coefficient: Coefficient estimate on the dummy (representing the effect in school systems without external exit exams). – Interaction: Coefficient estimate on the interaction term between the dummy and external exit exams (representing the difference in the effect between school systems without and with external exit exams). – Clustering-robust standard error in parentheses, using schools as level of clustering. – WLS regressions using students' sampling probabilities as weights. – Central exams [1], background [14], home inputs [6], resources [8], institutions [11], institutional interactions [11], imputation dummies [41]. – Significance levels (based on clustering-robust standard errors): * 1 percent. + 5 percent. ° 10 percent.

significant difference in any of the three subjects depending on whether central exams are in place. Second, there is also little evidence of effect heterogeneity by immigration status. The only statistically significant difference is that central exams dampen the effect of the country of birth of the mother in science, which is in line with the TIMSS results. Third, there is not much evidence of effect heterogeneity by family status, where statistically significant effects occur only in science, with the ambiguous result that both the single-father and the both-parents categories perform higher. Fourth, in terms of parental education, work status and occupation, there are no statistically significant effect differences in science. In math, the most notable difference is that the performance lag of children from blue-collar parents is larger in central-exam systems. Fifth, there are more statistically significant differences in reading performance with respect to parental education, work status and occupation, with a tendency of family-background differences to be larger in central-exam systems. This result is somewhat unexpected, given that the central-exam variable is the simple mean of math and science and thus a measure containing substantial measurement error.

In sum, the results point to relatively little heterogeneity by family background in the effect of central exams on PISA performance, while the disadvantage of coming from an immigrant or less-educated family background on TIMSS performance seems to be reduced by central exams, particularly in math.

4.2.3 Parental Interest

It was argued in Section 2.1.2 that the positive effect of central exams may be especially salient when parents are strongly concerned about their children's educational progress, but not as much when parents are less concerned. The TIMSS teacher background questionnaires contain two measures that can help to test for this differential impact. First, teachers reported to what extent, in their view, parents *uninterested* in their children's learning and progress limit how the teachers teach their class. Second, teachers also reported whether their teaching is limited by parents *interested* in their children's progress.

The results on the differential impact of these parental-influence measures with and without central exams are reported alongside the family-background results in Table 1. Students whose teachers reported that their teaching was substantially limited by uninterested parents performed worse than students whose teachers reported that their teaching was not limited by uninterested parents. This tendency was even larger in systems with central exams. In systems without central exams, students whose teachers reported that their teaching was limited a lot by interested parents again performed worse. However, in central-exam systems,

students whose teachers reported that interested parents limited how they teach their class performed just as well in math as students whose teachers did not say so, and in science, the effect even turns around into a statistically significant positive one. That is, even though teachers judged the intrusion of interested parents as limiting their teaching, student performance in fact did not suffer, but rather benefited from this “limitation.”

While the involvement of interested parents may limit student performance in systems without central exams because parents do not have well-founded information on which to base their interventions, central-exam systems seem to ensure that interested parents have the information necessary to intervene properly. Parents uninterested in their children’s educational progress do not seem to make use of this information, and their lack of interest hurts students’ educational performance. But it seems that the involvement of interested parents does not go all the way to being detrimental when central exams are in place, even when teachers might judge it to be so. In conclusion, the effect of central exams seems to be heterogeneous with respect to parental interest.

5. Evidence on Heterogeneity by School Characteristics

This section provides empirical evidence on the theoretical predictions presented in Section 2.2 about how the effect of central exams on student performance may differ between schools with different extents of autonomy.

5.1 The Interaction of Central Exams and School Autonomy

In testing the hypotheses on effect heterogeneity of central exams by school autonomy empirically, the interest again lies in testing effect heterogeneity along an observed dimension. Thus, interaction effects between central exams and the observed institutional features, which include measures of school autonomy, are simply added to equation (2) above:

$$(3) \quad T_{ilsc} = \alpha E_c + (E_c B_{ilsc}) \lambda + B_{ilsc} \beta + R_{lsc} \gamma + (E_c I_{lsc}) \theta + I_{lsc} \delta + a + \mu_c + \eta_s + \nu_l + \varepsilon_i \quad .$$

The estimated interaction effects show whether the effect of school autonomy in various decision-making areas differs between education systems with and without central exams. Or, put differently, they show whether the effect of central exams is heterogeneous with respect to the extent of school autonomy.

Table 3 reports the results on autonomy and interaction effects in TIMSS, and Table 4 in PISA. The following discussion focuses on the most striking findings with respect to effect

heterogeneity of central exams, which are also depicted in Figure 3 for the case of math. The diagrams in Figure 3 show, for different decision-making areas, student performance under the four conditions resulting from the presence and absence of school autonomy and central exams: the performance of students in schools without autonomy in systems without central exams; with autonomy but without central exams; without autonomy but with central exams; and with both autonomy and central exams. In each diagram, student performance is shown relative to the condition with the lowest performance.¹²

< Tables 3 and 4 and Figure 3 about here >

5.2 Decisions With Opportunism and With Local Knowledge Lead

Figures 3a and 3b show the case of whether schools are responsible for deciding on teacher salaries, one estimated using TIMSS data, the other using PISA data. In systems without central exams, school autonomy regarding teacher salaries has a *negative* effect on student performance. In systems with central exams, student performance is generally higher than in systems without central exams, both in cases with and without school autonomy. In addition, however, it is striking that the effect of school autonomy on student performance in systems with central exams is turned completely around. Salary autonomy of schools has *positive* effects on student performance in central-exam systems.¹³ That is, there is substantial heterogeneity in the effect of central exams on student performance depending on whether schools have autonomy over teacher salaries, with the effect being much stronger for autonomous schools.

Decisions on teacher salaries thus appear to involve both incentives for opportunistic behaviour and local knowledge leads (cf. Figure 1). Without central exams, the negative performance effect of opportunistic decisions taken by the schools dominates, as this local opportunistic behaviour cannot be externally observed and thus cannot be sanctioned. Hence school decision-makers do not feel obliged to set teacher salaries so as to contribute to enhancing student performance, but can use their decision-making autonomy to promote other interests. In contrast, central exams provide information about whether the schools perform well or not, so that supervisory authorities and parents can draw possible consequences from

¹² The estimates on which these diagrams are based control for all the control variables of family background, resources and institutions of equation (1), but – unlike the results reported in Tables 4 and 5 – not for further interaction effects between central exams and family or institution variables. Otherwise, the specific effect of central exams would be estimated quite imprecisely and the bars for the effects in central-exam systems would consequently be based on imprecise estimates (Wößmann, 2002).

¹³ In PISA, only the math results are statistically significant.

Table 3: Interaction Effects between Central Exams and Institutional Settings in TIMSS

	Math		Science	
	Coefficient	Interaction	Coefficient	Interaction
<i>School responsibility</i>				
School budget	-6.9 ⁺ (2.8)	7.7 ⁺ (3.5)	-12.0* (2.6)	16.1* (3.5)
Purchasing supplies	7.1 ⁺ (3.2)	-5.7 (5.0)	15.6* (3.1)	-6.2 (5.4)
Hiring teachers	21.6* (2.6)	-20.2* (3.1)	0.3 (1.9)	4.6 ^o (2.6)
Determining teacher salaries	-28.3* (3.6)	50.2* (4.1)	-8.2* (2.6)	29.2* (3.1)
<i>Teachers' influence</i>				
Class teacher has strong influence on				
Money for supplies	-24.7* (5.1)	29.1* (6.3)	-6.9 ^o (3.6)	13.6* (4.5)
Kind of supplies	3.0 (2.8)	-3.5 (3.8)	6.0* (2.0)	-3.7 (2.9)
Subject matter	-12.3* (2.3)	8.7* (2.8)	-4.6* (1.7)	-0.7 (2.2)
Textbook	11.6* (3.1)	-11.7* (3.6)	6.3* (1.8)	-9.9* (2.6)
Strong influence on curriculum				
Teacher individually	14.6* (2.1)	-3.9 (2.7)	14.5* (1.8)	-7.4* (2.5)
Subject teachers	-5.0 ⁺ (2.4)	2.8 (3.1)	-5.8* (2.1)	8.2* (2.8)
School teachers collectively	-14.7* (2.1)	6.5 ⁺ (2.8)	-15.3* (1.9)	14.4* (2.6)
Teacher unions	-8.5 (5.4)	-29.5* (8.7)	-6.7 (5.1)	-30.0* (9.1)
Students (observations)	447,089		447,089	
Schools (PSUs)	12,175		12,175	
Countries	77		77	
R ²	0.296		0.266	

Dependent variable: TIMSS international math/science test score (pooled TIMSS-95 and TIMSS-Repeat). – Coefficient: Coefficient estimate on the dummy (representing the effect in school systems without external exit exams). – Interaction: Coefficient estimate on the interaction term between the dummy and external exit exams (representing the difference in the effect between school systems without and with external exit exams). – Clustering-robust standard error in parentheses, using schools as level of clustering. – WLS regressions using students' sampling probabilities as weights. – Controls [number of variables]: Central exams [1], background [22], background interactions [18], resources [13], centralisation and student incentives [6]. – Significance levels (based on clustering-robust standard errors): * 1 percent. – ⁺ 5 percent. – ^o 10 percent.

Source: Based on Wößmann (2002).

Table 4: Interaction Effects between Central Exams and Institutional Settings in PISA

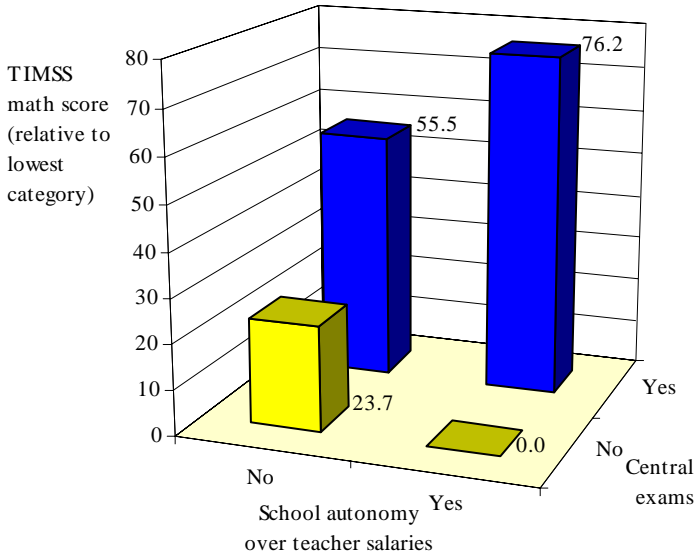
	Math		Science		Reading	
	Coefficient	Interaction	Coefficient	Interaction	Coefficient	Interaction
<i>School autonomy</i>						
Determining course content	-6.2 ⁺ (2.5)	19.8* (3.9)	-4.8 ⁺ (2.1)	14.3* (3.9)	-8.6* (2.3)	19.6* (4.0)
Choosing textbooks	-2.0 (5.3)	48.3* (10.1)	2.0 (4.1)	58.7* (9.2)	0.1 (4.9)	61.0* (10.1)
Formulating school budget	-4.2 (2.8)	-5.9 (4.1)	-5.2 ⁺ (2.5)	-1.3 (3.9)	-0.1 (2.7)	-4.9 (4.1)
Deciding on budget allocations within school	5.5 (4.7)	11.6 ^o (6.9)	11.4* (4.4)	5.6 (6.5)	7.5 (4.6)	4.3 (6.3)
Hiring teachers	16.5* (3.4)	-12.7 ⁺ (5.0)	2.7 (3.1)	-5.4 (5.0)	7.8 ⁺ (3.2)	-8.5 ^o (5.1)
Firing teachers	-4.2 (3.6)	4.8 (4.7)	-5.6 ^o (3.3)	6.6 (4.6)	-6.7 ⁺ (3.3)	2.7 (4.7)
Establishing teachers' starting salaries	-19.4* (6.6)	23.3* (8.0)	-5.1 (5.0)	3.4 (6.1)	-5.7 (5.6)	0.6 (6.8)
Determining teachers' salary increases	1.6 (6.4)	-1.9 (7.7)	-3.6 (4.7)	2.2 (5.9)	-1.0 (5.3)	0.7 (6.6)
<i>Public vs. private operation and funding</i>						
Publicly managed school	-16.7* (3.5)	1.2 (4.9)	-12.6* (2.7)	-0.4 (4.0)	-16.5* (2.8)	5.3 (4.1)
Government funding (share)	8.2 ^o (4.9)	-0.3 (7.6)	-1.9 (4.0)	4.1 (6.5)	1.1 (4.1)	-1.7 (6.9)
<i>Regularity of testing</i>						
Standardised tests	-6.4* (2.4)	11.7* (3.6)	-8.5* (2.1)	15.8* (3.3)	-6.3* (2.2)	13.6* (3.3)
Students (observations)	96,855		96,758		174,227	
Schools (PSUs)	6,611		6,613		6,626	
Countries	31		31		31	
R ²	0.331		0.285		0.335	

Dependent variable: PISA international math/science/reading test score. – Coefficient: Coefficient estimate on the dummy (representing the effect in school systems without external exit exams). – Interaction: Coefficient estimate on the interaction term between the dummy and external exit exams (representing the difference in the effect between school systems without and with external exit exams). – Clustering-robust standard error in parentheses, using schools as level of clustering. – WLS regressions using students' sampling probabilities as weights. – Controls [number of variables]: Central exams [1], background [36], home inputs [6], resources [8], imputation dummies [41]. – Significance levels (based on clustering-robust standard errors): * 1 percent. – ⁺ 5 percent. – ^o 10 percent.

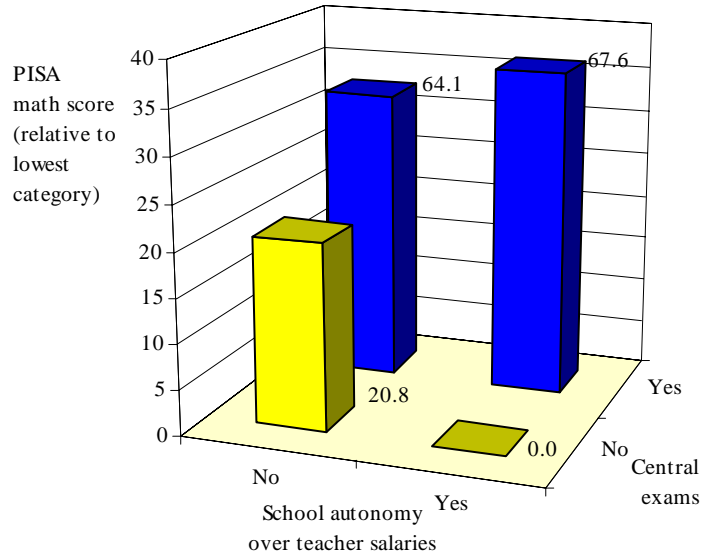
Source: Fuchs and Wößmann (2004).

Figure 3: Interaction Effects between Central Exams and School Autonomy

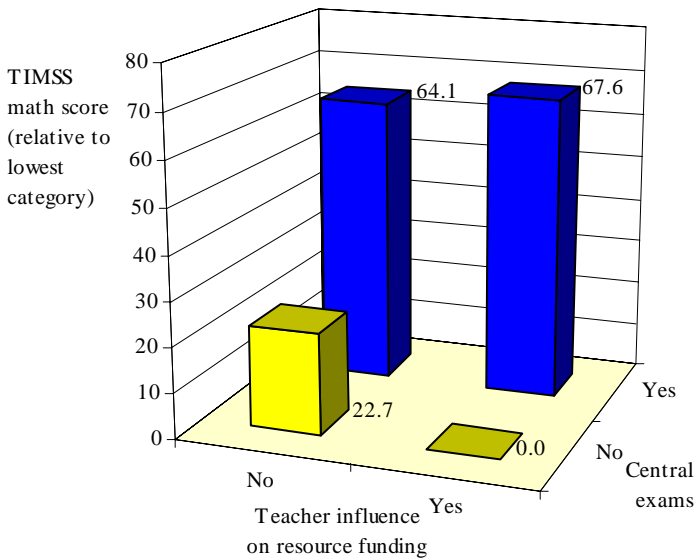
(a) School autonomy over teacher salaries in TIMSS and TIMSS-Repeat



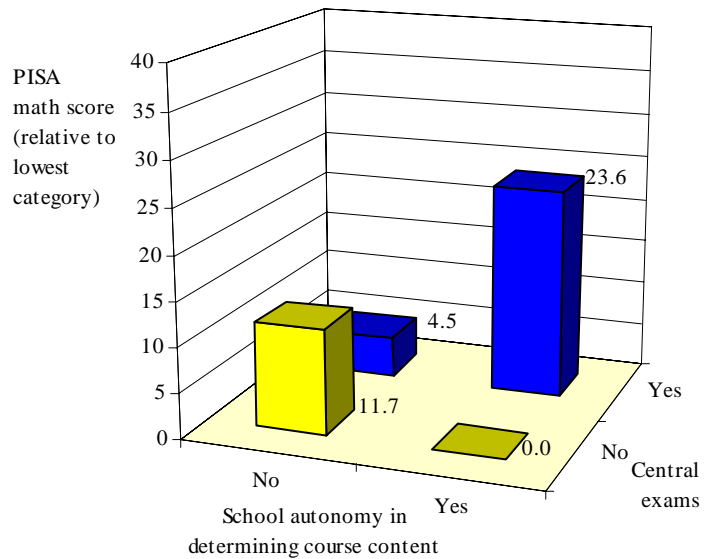
(b) School autonomy over teacher salaries in PISA



(c) Teacher influence on resource funding in TIMSS and TIMSS-Repeat



(d) School autonomy in determining course content in PISA



poor school behaviour. This creates incentives for decision-makers in schools not to exploit their autonomy in setting teacher salaries in an opportunistic way, but to use it in order to effectively promote student performance. The benefits of superior local knowledge then come into effect, as school decision-makers ought to know better than any central authority which teachers deserve to be rewarded for good work. The size of the performance difference of students in schools with salary autonomy between systems with and without central exams is roughly as large as two grade-level equivalents in the TIMSS studies.

The case is similar when decisions on school resources are decentralised in such a way that teachers have a say in the money available for purchasing supplies, based on data from the TIMSS studies which feature extensive teacher background questionnaires (Figure 3c). In this decision-making area too, decentralised decision-making autonomy has a negative effect on student performance in systems without central exams, whereas it has a positive effect in systems with central exams. However, the difference between schools with and without teacher influence on resource funding in systems with central exams is not statistically significant. This could be due either to the fact that opportunistic behaviour is not entirely prevented by central exams in this case and consequently weakens the positive effect of local knowledge, or to the fact that there is no significant local knowledge lead.

A similar pattern can also be observed for school autonomy in determining course contents, reported in the PISA study (Figure 3d). In systems without external exit exams, students in schools that have autonomy in determining course contents perform statistically significantly worse than otherwise. That is, the effect of school autonomy in this area seems to be negative if there are no external exit exams to hold schools accountable for what they are doing. Again, the effect turns around to be statistically significantly positive if schools are made accountable for their behaviour through external exit exams. This pattern of results suggests that the decision-making area of determining course contents also entails substantial incentives for local opportunistic behaviour as well as significant local knowledge lead. The incentives for local opportunistic behaviour stem from the fact that content decisions influence the workload of the teachers, and they lead to the negative autonomy effect in systems without accountability. The local knowledge lead stems from the fact that teachers probably know best what specific course contents would be best suited for their specific students, and it leads to the positive autonomy effect in systems where external exit exams mitigate the scope for opportunistic behaviour.

5.3 Decisions With Opportunism and Without Local Knowledge Lead

There are other dimensions of school autonomy along which the effect of central exams differs, albeit not as strongly as in the previous cases. Thus, in systems without central exams, school autonomy in budgeting has a negative impact on TIMSS performance (Table 3), which may be due to incentives for opportunistic behaviour in funding. In systems with central exams, this negative effect of school autonomy disappears, although without turning into a significant positive effect. This could be due to the fact that external agencies need by no means have a knowledge disadvantage in budget questions compared to individual schools which often lack the required specialist staff.

The same argument might be raised for the task of determining the subject matter to be covered in class, which shows a similar pattern depending on how much autonomy teachers have in the task (Table 3). A similar tendency also emerges for school autonomy in decisions on firing teachers, reported only in PISA, in all three subjects (Table 4). The relationship between firing autonomy and science and reading performance is statistically significantly negative in systems without external exit exams, but not in systems with external exit exams.

The same pattern also emerges depending on whether teachers of a school collectively have a say in the curriculum to be taught, reported only in TIMSS (Table 3). Without monitoring by central exams, collective teacher influence has a negative impact on student performance, which may be due to opportunistic interests of teachers as regards the workload to be fulfilled. If a central-exam system does exist, this negative performance effect is attenuated into an insignificant effect of teacher autonomy. This can be rationalised within the framework of the above model by assuming attenuated opportunistic behaviour with the simultaneous absence of local knowledge leads where decisions are taken collectively by the teachers.¹⁴

5.4 Further Decision-Making Areas

Things look different when individual teachers rather than the teachers' collective can influence the curriculum (Table 3). In this case, a positive effect of teacher autonomy on student performance is observed in systems with and without central exams, showing no statistically significant difference between the two systems in math (and a slightly attenuated effect in science). In the model framework considered here, this would mean that the

¹⁴ Note that the underlying estimate controls for the influence of individual teachers on the curriculum, so that the individual knowledge benefits of the teachers are held constant.

individual teachers are unable to push through opportunistic behaviours in addition to their collective influence, but that they possess local knowledge advantages as individuals.

In PISA, a positive relationship between school autonomy in deciding on budget allocations within schools (within the framework of a formulated overall school budget) tends to be stronger in systems with external exit exams (Table 4). The difference between the two kinds of systems is statistically significant only in math, though. The pattern suggests that this decision-making area features only small incentives for local opportunistic behaviour, but a significant local knowledge lead.

In systems without external exit exams, there is no statistically significant relationship between school autonomy in choosing textbooks and student achievement in either subject on the PISA test (Table 4). However, there is a substantial statistically significant positive relationship in systems with external exit exams. This pattern reflects the theoretical case where incentives for local opportunistic behaviour are offset by a local knowledge lead. External exit exams suppress the negative opportunism effect and keep the positive knowledge-lead effect. Thus, the beneficial effects of school autonomy in textbook choice prevail only in systems where schools are made accountable for their behaviour through external exit exams.¹⁵

No statistically significant relationship is detected between school autonomy in hiring teachers and student performance in science in PISA, and in science in TIMSS, the relationship turns slightly positive in central-exam systems. In math and reading, however, the relationship between hiring autonomy and student performance is statistically significantly positive in systems without external exit exams, but not in systems with external exit exams. It seems hard to rationalise this result in the framework of the theoretical model. It may be that the pattern reflects a positive selection effect of teacher choice on the detriment of non-autonomous schools in systems without external exit exams, which might be less pronounced in the more transparent external-exam systems.

PISA also provides school-level data on the public versus private management and funding of schools (cf. Fuchs and Wößmann, 2004, for details). The results are found to be independent of whether the school system has external exit exams are not (Table 4). In both

¹⁵ TIMSS reports whether teachers have a lot of influence on textbooks, and the results are quite different (Table 3). If individual teachers in systems without central exams have a say in the textbooks used, this has a positive effect on student performance. This is likely to be due to local knowledge leads which are not counteracted by opportunistic interests, as the teachers would do themselves a disservice if they were to select unsuitable books. This positive effect of teacher autonomy is smaller and statistically insignificant in systems with central exams (cf. Wößmann, 2003c).

kinds of systems, students perform statistically significantly better in privately managed schools. Also in both kinds of systems, science and reading performance does not vary with schools' share of public funding, while math performance increases with public funding.

The general pattern of results strongly suggests that there is effect heterogeneity of central exams with respect to school characteristics, particularly in terms of the autonomy of schools and teachers. External exit exams and school autonomy are complementary institutional features of a school system in any decision-making area that includes scope for opportunism and local knowledge leads. School autonomy tends to be more beneficial for student performance in all subjects when external exit exams are in place to hold the autonomous schools accountable for their decision-making. This evidence corroborates the reasoning of external exams as the “currencies” of school systems (Wößmann, 2003c) which ensure that otherwise decentralised systems function in the interest of students' educational performance.

6. Evidence on Heterogeneity by Time

This section presents empirical evidence on how the effect of central exams on student performance differs along the time dimension (cf. Section 2.3 for the theoretical background). First, it looks at whether the effect of central exit exams at the end of secondary schools increases during secondary school. Second, it estimates whether regular standardised testing throughout secondary school exerts effects in addition to the effect of central exit exams.

6.1 The Effect of Central Exit Exams by Grade

The first theoretical prediction of effect heterogeneity over time was that the effect of central exit exams may increase with grade levels. This effect heterogeneity of central exit exams can be tested using the TIMSS-95 data, which allows a comparison between seventh- and eighth-grade performance.

The results on the interaction effect between central exams and grade level are included alongside the family-background results in Table 1. The impact of central exams on TIMSS math performance is 17.5 percent of a standard deviation larger in eighth grade than in seventh grade. Likewise, the impact of central exams on eighth-grade science performance is 12.2 percent of a standard deviation larger than their impact in seventh grade. Thus, the impact that central exit exams exert on student performance grows over the course of secondary education, showing effect heterogeneity over time.

6.2 The Effect of Regular Standardised Testing

The second theoretical prediction of effect heterogeneity over time was that regular standardised testing during the course of schooling may have positive effects that add to the effects of central school-leaving exams. The PISA study provides school-level information on whether standardised testing was used for fifteen-year-old students at least once a year, which allows for a test of this hypothesis.

As the results reported in Table 4 show, the relationship between standardised tests and student achievement differs strongly and statistically significantly between systems with and without external exit exams. If there are no external exit exams, standardised testing is statistically significantly negatively related to student achievement in all three subjects. One way to explain this finding would be that if the educational goals and standards of the school system are not clearly specified, standardised testing can backfire and lead to weaker student performance. But the relationship between standardised testing and student achievement in all three subjects turns around to be statistically significantly positive in systems where external exit exams are in place.

Figure 4 depicts this complementary relationship between central exit exams and earlier standardised testing,¹⁶ which shows a pattern similar to the effect heterogeneity with respect to school autonomy in decisions that involve both opportunism and local knowledge leads (cf. Section 5.2). That is, regular standardised examination seems to have additional positive performance effects when added to central exit exams. Thus, there is an additional effect heterogeneity of standardised exams along the time dimension, in that the effect of central exit exams at the end of secondary schooling can be supplemented by earlier standardised testing.

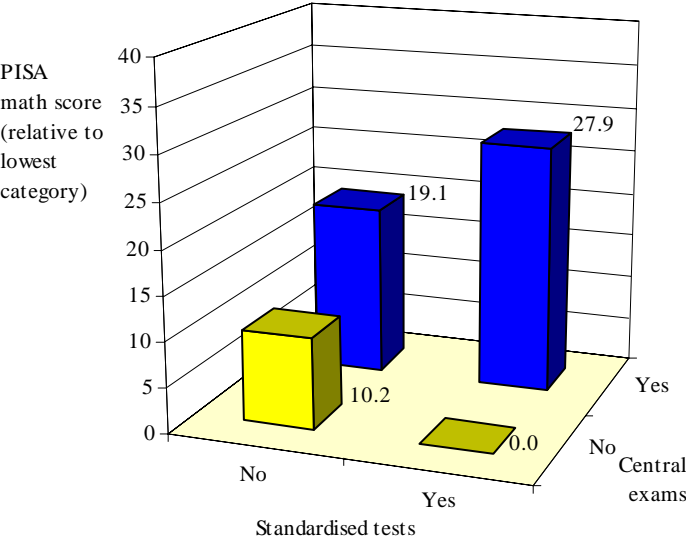
< Figure 4 about here >

7. Summary and Conclusions

The paper has analysed the heterogeneity of the effect of central exams on student performance along three dimensions, presenting evidence from three international student achievement tests (TIMSS, TIMSS-Repeat and PISA). First, the paper has looked at differences in the central-exam effect by student characteristics. It was argued that due to the lower labour-market mobility of less-skilled workers and a reasonable capability of local

¹⁶ As in Figure 3, the regression on which Figure 4 is based controls for family, resource and institutional controls, but not for interaction terms between central exams and other variables.

Figure 4: Central Exit Exams and Regular Standardised Testing



employers to assess local grades, the effect of central exams on student performance may be lower for low-ability students relative to high-ability students. Quantile-regression estimates support this pattern for most of the international tests. By including interaction effects between family-background measures and central exams, it could be shown that the effect of central exams does not vary substantially along most family-background dimensions. The main heterogeneity is that in TIMSS, the disadvantage of coming from an immigrant or less-educated family background seems to be reduced by central exams. More evidently, parental involvement gets better informed in central-exam systems, so that the interest that parents bring to the teaching of their children is more beneficial when central exams are in place.

Second, the paper has looked at heterogeneity in the effect of central exams by school and teacher autonomy. In a principal-agent framework of educational production, it could be shown that the effect of central exams depends on the extent of school autonomy, particularly in decision-making areas that contain incentives for opportunistic behaviour. In such areas, central exams can avert the disadvantages of opportunistic behaviour by introducing accountability, bringing the advantages of local knowledge to the fore. The evidence shows that there is indeed substantial heterogeneity in the effect of central exams depending on whether schools have autonomy. The effect of central exams is much stronger for schools that have autonomy in such decision-making areas as establishing teacher salaries, determining course contents, budgeting, choosing textbooks, determining the subject matter to be covered in class and for schools whose class teachers can influence resource funding. Thus, there is a complementarity between central exams and school autonomy.

Third, the paper has analysed whether the effect of external testing may vary over the course of secondary education. It was found that the effect of central exit exams at the end of secondary schools on how much students learn during secondary school increases from seventh to eighth grade. Also, regular standardised examinations during secondary education exert additional positive performance effects when added to central exit exams.

In sum, there is clear evidence that the performance effects of central exams are heterogeneous, differing along the student, school and time dimensions. The evidence not only shows who gains most from central exams, but also that central exit exams are particularly performance-conducive once combined with school autonomy and regular external testing.

References

- Adams, R. & Wu, M. (Eds.) (2002) *PISA 2000 Technical Report* (Paris, OECD).
- Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.L. & Smith T.A. (1996) *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS)* (Chestnut Hill, MA, Boston College).
- Bishop, J.H. (1995) The Impact of Curriculum-Based External Examinations on School Priorities and Student Learning, *International Journal of Educational Research*, 23, pp. 653-752.
- Bishop, J.H. (1997) The Effect of National Standards and Curriculum-Based Exams on Achievement, *American Economic Review*, 87, pp. 260-264.
- Bishop, J.H. (1999) Are National Exit Examinations Important for Educational Efficiency?, *Swedish Economic Policy Review*, 6, pp. 349-398.
- Bishop, J.H. (2004) Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn, forthcoming in: Hanushek, E.A. & Welch, F. (Eds.) *Handbook of the Economics of Education* (Amsterdam, North-Holland).
- Bishop, J.H., Mane, F., Bishop, M. & Moriarty, J.Y. (2001) The Role of End-of-Course Exams and Minimum Competency Exams in Standards-Based Reforms, *Brookings Papers on Education Policy*, 2001, pp. 267-345.
- Bishop, J.H., Moriarty, J.Y. & Mane, F. (2000) Diplomas for Learning, Not Seat Time: The Impact of New York Regents Examinations, *Economics of Education Review*, 19, pp. 333-349.
- Bishop, J.H. & Wößmann, L. (2004) Institutional Effects in a Simple Model of Educational Production, *Education Economics*, 12, pp. 17-38.
- Deaton, A. (1997) *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy* (Baltimore, The Johns Hopkins University Press).
- DuMouchel, W.H. & Duncan, G.J. (1983) Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples, *Journal of the American Statistical Association*, 78, pp. 535-543.
- Eide, E. & Showalter, M.H. (1998) The Effect of School Quality on Student Performance: A Quantile Regression Approach, *Economics Letters*, 58, pp. 345-350.
- Fertig, M. (2003) Who's to Blame? The Determinants of German Students' Achievement in the PISA 2000 Study, IZA Discussion Paper 739 (Bonn, Institute for the Study of Labor).
- Fuchs, T. & Wößmann, L. (2004) What Accounts for International Differences in Student Performance? A Re-examination using PISA Data, CESifo Working Paper 1235 (Munich, CESifo).
- Gonzalez, E.J. & Miles, J.A. (Eds.) (2001) *User Guide for the TIMSS 1999 International Database* (Chestnut Hill, MA, Boston College).
- Gonzalez, E.J. & Smith, T.A. (Eds.) (1997) *User Guide for the TIMSS International Database: Primary and Middle School Years* (Chestnut Hill, MA, Boston College).
- Hanushek, E.A. (2002) Publicly Provided Education, in: Auerbach, A.J. & Feldstein, M. (Eds.) *Handbook of Public Economics, Vol. 4*, pp. 2045-2141 (Amsterdam, North-Holland).

- Jürges, H., Schneider, K. & Büchel, F. (2003) The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany, CESifo Working Paper 939 (Munich, CESifo).
- Koenker, R. & Bassett, G.Jr. (1978) Regression Quantiles, *Econometrica*, 46, pp. 33-50.
- Laffont, J.-J. & Martimort, D. (2002) *The Theory of Incentives: The Principal-Agent Model* (Princeton, Princeton University Press).
- Levin, J. (2001) For Whom the Reductions Count: A Quantile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement, *Empirical Economics*, 26, pp. 221-246.
- Martin, M.O. & Kelly, D.L. (Eds.) (1996) *Third International Mathematics and Science Study Technical Report, Vol. I: Design and Development* (Chestnut Hill, MA, Boston College).
- Martin, M.O. & Kelly, D.L. (Eds.) (1997) *Third International Mathematics and Science Study Technical Report, Vol. II: Implementation and Analysis – Primary and Middle School Years* (Chestnut Hill, MA, Boston College).
- Martin, M.O., Gregory, K.D. & Stemler, S.E. (Eds.) (2000) *TIMSS 1999 Technical Report* (Chestnut Hill, MA, Boston College).
- Moulton, B.R. (1986) Random Group Effects and the Precision of Regression Estimates, *Journal of Econometrics*, 32, pp. 385-397.
- Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Gregory, K.D., Garden, R.A., O'Connor, K.M., Chrostowski, S.J. & Smith, T.A. (2000) *TIMSS 1999 International Mathematics Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade* (Chestnut Hill, MA, Boston College).
- Organisation for Economic Co-operation and Development (OECD) (2000) *Measuring Student Knowledge and Skills: The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy* (Paris, OECD).
- Organisation for Economic Co-operation and Development (OECD) (2001) *Knowledge and Skills for Life: First Results from the OECD Programme for International Student Assessment (PISA) 2000* (Paris, OECD).
- Organisation for Economic Co-operation and Development (OECD) (2002) *Manual for the PISA 2000 Database* (Paris, OECD).
- Wooldridge, J.M. (2001) Asymptotic Properties of Weighted *M*-Estimators for Standard Stratified Samples, *Econometric Theory*, 17, pp. 451-470.
- Wößmann, L. (2002) *Central Examinations Improve Educational Performance: International Evidence*, Kiel Discussion Papers 397 (Kiel, Institute for World Economics).
- Wößmann, L. (2003a) Schooling Resources, Educational Institutions and Student Performance: the International Evidence, *Oxford Bulletin of Economics and Statistics*, 65, pp. 117-170.
- Wößmann, L. (2003b) Central Exit Exams and Student Achievement: International Evidence, in: Peterson, P.E. & West, M.R. (Eds.) *No Child Left Behind? The Politics and Practice of School Accountability*, pp. 292-323 (Washington, D.C., Brookings Institution Press).
- Wößmann, L. (2003c) Central Exams as the “Currency” of School Systems: International Evidence on the Complementarity of School Autonomy and Central Exams, *DICE Report - Journal for Institutional Comparisons*, 1, pp. 46-56.
- Wößmann, L. (2004) How Equal Are Educational Opportunities? Family Background and Student Achievement in Europe and the United States, CESifo Working Paper 1162 (Munich, CESifo).

CESifo Working Paper Series

(for full list see www.cesifo.de)

- 1269 Thomas Eichner and Rüdiger Pethig, Economic Land Use, Ecosystem Services and Microfounded Species Dynamics, September 2004
- 1270 Federico Revelli, Performance Rating and Yardstick Competition in Social Service Provision, September 2004
- 1271 Gerhard O. Orosel and Klaus G. Zauner, Vertical Product Differentiation When Quality is Unobservable to Buyers, September 2004
- 1272 Christoph Böhringer, Stefan Boeters, and Michael Feil, Taxation and Unemployment: An Applied General Equilibrium Approach, September 2004
- 1273 Assaf Razin and Efraim Sadka, Welfare Migration: Is the Net Fiscal Burden a Good Measure of its Economics Impact on the Welfare of the Native-Born Population?, September 2004
- 1274 Tomer Blumkin and Volker Grossmann, Ideological Polarization, Sticky Information, and Policy Reforms, September 2004
- 1275 Katherine Baicker and Nora Gordon, The Effect of Mandated State Education Spending on Total Local Resources, September 2004
- 1276 Gabriel J. Felbermayr and Wilhelm Kohler, Exploring the Intensive and Extensive Margins of World Trade, September 2004
- 1277 John Burbidge, Katherine Cuff and John Leach, Capital Tax Competition with Heterogeneous Firms and Agglomeration Effects, September 2004
- 1278 Joern-Steffen Pischke, Labor Market Institutions, Wages and Investment, September 2004
- 1279 Josef Falkinger and Volker Grossmann, Institutions and Development: The Interaction between Trade Regime and Political System, September 2004
- 1280 Paolo Surico, Inflation Targeting and Nonlinear Policy Rules: The Case of Asymmetric Preferences, September 2004
- 1281 Ayal Kimhi, Growth, Inequality and Labor Markets in LDCs: A Survey, September 2004
- 1282 Robert Dur and Amihai Glazer, Optimal Incentive Contracts for a Worker who Envis his Boss, September 2004
- 1283 Klaus Abberger, Nonparametric Regression and the Detection of Turning Points in the Ifo Business Climate, September 2004

- 1284 Werner Güth and Rupert Sausgruber, Tax Morale and Optimal Taxation, September 2004
- 1285 Luis H. R. Alvarez and Erkki Koskela, Does Risk Aversion Accelerate Optimal Forest Rotation under Uncertainty?, September 2004
- 1286 Giorgio Brunello and Maria De Paola, Market Failures and the Under-Provision of Training, September 2004
- 1287 Sanjeev Goyal, Marco van der Leij and José Luis Moraga-González, Economics: An Emerging Small World?, September 2004
- 1288 Sandro Maffei, Nikolai Raabe and Heinrich W. Ursprung, Political Repression and Child Labor: Theory and Empirical Evidence, September 2004
- 1289 Georg Götz and Klaus Gugler, Market Concentration and Product Variety under Spatial Competition: Evidence from Retail Gasoline, September 2004
- 1290 Jonathan Temple and Ludger Wößmann, Dualism and Cross-Country Growth Regressions, September 2004
- 1291 Ravi Kanbur, Jukka Pirttilä and Matti Tuomala, Non-Welfarist Optimal Taxation and Behavioral Public Economics, October 2004
- 1292 Maarten C. W. Janssen, José Luis Moraga-González and Matthijs R. Wildenbeest, Consumer Search and Oligopolistic Pricing: An Empirical Investigation, October 2004
- 1293 Kira Börner and Christa Hainz, The Political Economy of Corruption and the Role of Financial Institutions, October 2004
- 1294 Christoph A. Schaltegger and Lars P. Feld, Do Large Cabinets Favor Large Governments? Evidence from Swiss Sub-Federal Jurisdictions, October 2004
- 1295 Marc-Andreas Müндler, The Existence of Informationally Efficient Markets When Individuals Are Rational, October 2004
- 1296 Hendrik Jürges, Wolfram F. Richter and Kerstin Schneider, Teacher Quality and Incentives: Theoretical and Empirical Effects of Standards on Teacher Quality, October 2004
- 1297 David S. Evans and Michael Salinger, An Empirical Analysis of Bundling and Tying: Over-the-Counter Pain Relief and Cold Medicines, October 2004
- 1298 Gershon Ben-Shakhar, Gary Bornstein, Astrid Hopfensitz and Frans van Winden, Reciprocity and Emotions: Arousal, Self-Reports, and Expectations, October 2004
- 1299 B. Zorina Khan and Kenneth L. Sokoloff, Institutions and Technological Innovation During Early Economic Growth: Evidence from the Great Inventors of the United States, 1790 – 1930, October 2004

- 1300 Piero Gottardi and Roberto Serrano, Market Power and Information Revelation in Dynamic Trading, October 2004
- 1301 Alan V. Deardorff, Who Makes the Rules of Globalization?, October 2004
- 1302 Sheilagh Ogilvie, The Use and Abuse of Trust: Social Capital and its Deployment by Early Modern Guilds, October 2004
- 1303 Mario Jametti and Thomas von Ungern-Sternberg, Disaster Insurance or a Disastrous Insurance – Natural Disaster Insurance in France, October 2004
- 1304 Pieter A. Gautier and José Luis Moraga-González, Strategic Wage Setting and Coordination Frictions with Multiple Applications, October 2004
- 1305 Julia Darby, Anton Muscatelli and Graeme Roy, Fiscal Federalism, Fiscal Consolidations and Cuts in Central Government Grants: Evidence from an Event Study, October 2004
- 1306 Michael Waldman, Antitrust Perspectives for Durable-Goods Markets, October 2004
- 1307 Josef Honerkamp, Stefan Moog and Bernd Raffelhüschen, Earlier or Later: A General Equilibrium Analysis of Bringing Forward an Already Announced Tax Reform, October 2004
- 1308 M. Hashem Pesaran, A Pair-Wise Approach to Testing for Output and Growth Convergence, October 2004
- 1309 John Bishop and Ferran Mane, Educational Reform and Disadvantaged Students: Are They Better Off or Worse Off?, October 2004
- 1310 Alfredo Schclarek, Consumption and Keynesian Fiscal Policy, October 2004
- 1311 Wolfram F. Richter, Efficiency Effects of Tax Deductions for Work-Related Expenses, October 2004
- 1312 Franco Mariuzzo, Patrick Paul Walsh and Ciara Whelan, EU Merger Control in Differentiated Product Industries, October 2004
- 1313 Kurt Schmidheiny, Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland, October 2004
- 1314 David S. Evans, Andrei Hagiu and Richard Schmalensee, A Survey of the Economic Role of Software Platforms in Computer-Based Industries, October 2004
- 1315 Frank Riedel and Elmar Wolfstetter, Immediate Demand Reduction in Simultaneous Ascending Bid Auctions, October 2004
- 1316 Patricia Crifo and Jean-Louis Rullière, Incentives and Anonymity Principle: Crowding Out Toward Users, October 2004

- 1317 Attila Ambrus and Rossella Argenziano, Network Markets and Consumers Coordination, October 2004
- 1318 Margarita Katsimi and Thomas Moutos, Monopoly, Inequality and Redistribution Via the Public Provision of Private Goods, October 2004
- 1319 Jens Josephson and Karl Wärneryd, Long-Run Selection and the Work Ethic, October 2004
- 1320 Jan K. Brueckner and Oleg Smirnov, Workings of the Melting Pot: Social Networks and the Evolution of Population Attributes, October 2004
- 1321 Thomas Fuchs and Ludger Wößmann, Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and Use of Computers at Home and at School, November 2004
- 1322 Alberto Bisin, Piero Gottardi and Adriano A. Rampini, Managerial Hedging and Portfolio Monitoring, November 2004
- 1323 Cecilia García-Peñalosa and Jean-François Wen, Redistribution and Occupational Choice in a Schumpeterian Growth Model, November 2004
- 1324 William Martin and Robert Rowthorn, Will Stability Last?, November 2004
- 1325 Jianpei Li and Elmar Wolfstetter, Partnership Dissolution, Complementarity, and Investment Incentives, November 2004
- 1326 Hans Fehr, Sabine Jokisch and Laurence J. Kotlikoff, Fertility, Mortality, and the Developed World's Demographic Transition, November 2004
- 1327 Adam Elbourne and Jakob de Haan, Asymmetric Monetary Transmission in EMU: The Robustness of VAR Conclusions and Cecchetti's Legal Family Theory, November 2004
- 1328 Karel-Jan Alsem, Steven Brakman, Lex Hoogduin and Gerard Kuper, The Impact of Newspapers on Consumer Confidence: Does Spin Bias Exist?, November 2004
- 1329 Chiona Balfoussia and Mike Wickens, Macroeconomic Sources of Risk in the Term Structure, November 2004
- 1330 Ludger Wößmann, The Effect Heterogeneity of Central Exams: Evidence from TIMSS, TIMSS-Repeat and PISA, November 2004