DISCUSSION PAPER SERIES

IZA DP No. 2880

Fundamental Determinants of School Efficiency and Equity: German States as a Microcosm for OECD Countries

Ludger Wößmann

June 2007

Fundamental Determinants of School Efficiency and Equity: German States as a Microcosm for OECD Countries

Ludger Wößmann
University of Munich, Ifo Institute,
CESifo and IZA

Discussion Paper No. 2880
June 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 2880
June 2007

ABSTRACT

Fundamental Determinants of School Efficiency and Equity: German States as a Microcosm for OECD Countries*

Cross-country evidence on student achievement might be hampered by omitted country characteristics such as language or legal differences. This paper uses cross-state variation in Germany, whose sixteen states share the same language and legal system, but pursue different education policies. The same results found previously across countries hold within Germany: Higher mean student performance is associated with central exams, private school operation, and socio-economic background, but not with spending, while higher equality of opportunity is associated with reduced tracking. In a model that pools German states with OECD countries, these fundamental determinants do not differ significantly between the two samples.

JEL Classification: I28, L38, L33, H52, D02, D63, J24

Keywords: student performance, PISA, Germany, education production function, institutional effects in schooling

Corresponding author:

Ludger Woessmann
CESifo
Poschingerstr. 5
81679 Munich
Germany
E-mail: woessmann@ifo.de

* This research was performed during my visit at the Program on Education Policy and Governance at Harvard University, whose hospitality I gratefully acknowledge. Thanks to Rick Hanushek for pushing me to add the pooled Germany-OECD analysis and to Andreas Ammermüller, Sascha Becker, Elke Lüdemann, and Gabriela Schütz for comments.

1. Introduction

Cross-country evidence using international student achievement tests shows that institutional structures of school systems are important for the efficiency and equity of educational outcomes. In terms of efficiency, the main patterns of results suggest that students perform better in countries with external exit exams, school autonomy in process and personnel decisions, and publicly funded but privately operated schools.[1] While student performance is also strongly associated with students' socio-economic background, there is not much evidence of significant associations with schools' resource endowments. In terms of equity, the extent to which students' outcomes depend on their family background tends to be lower in countries with less intensive tracking, more extensive early childhood education, and publicly funded private operation.[2]

The virtue of the cross-country evidence is that it allows for an analysis of variation in institutional features which usually do not vary within countries. A key shortcoming, as with any international evidence, is that one can never be fully sure whether the observed features are exogenous to other country characteristics, i.e. whether some unobserved country heterogeneity remains. In an attempt to avoid bias from omitted variables, the literature has reverted to extensive sets of control variables for student, school, and country background. Specific studies have also made use of specifications with regional (continental) fixed effects in the estimation of effects of central exams (Wößmann 2003b) and of differences-in-differences specifications across grades in the estimation of effects of early tracking (Hanushek and Wößmann 2006).

This paper presents another solution to the problem of unobserved country heterogeneity. It analyzes variation within a country that has different school systems, namely Germany, where – due to the federal structure – the 16 states (Bundesländer) have essentially sole responsibility for education policy. The advantage of this approach is that the German systems provide variation in institutional features of the school system within the same country, same language, same legal system, same general culture, and the like.[3] On the other hand, a key drawback of the approach is that the number of relevant observations is only 16, leaving limited degrees of freedom and thus small variation to identify from. In addition, German confidentiality requirements preclude the

[1] See, e.g., Bishop (1997), Wößmann (2002, 2003a, 2003b), and Fuchs and Wößmann (2007). Wößmann (2007a, 2007b) reviews the international evidence on effects of institutions and resources on student performance.

[2] See Hanushek and Wößmann (2006), Schütz et al. (2005), and Ammermüller (2005).

[3] Note that apart from education policy, there is hardly any substantive policy area in which sole responsibility remains at the level of states in Germany.

informative use of student-level data,[4] so that this paper uses data at the state level only.[5] Despite the limited statistical power due to the small number of observations, the state sample should still allow for an identification of effects that are of first-order magnitude. At the same time, second-order effects may be hard to identify with substantial statistical confidence. However, by pooling German states with aggregate countries, we can test for differences in results between within-German and international country variation in a sample of up to 54 observations.

The German states differ dramatically in the efficiency and equity of schooling outcomes. When merging the 16 German states into one league table with the other 28 OECD countries that report PISA 2003 performance data, Bavaria (the second largest state in terms of students) takes 5th place internationally (test score 533), while North Rhine-Westphalia (the largest state) takes 35th place (486) and the city-state of Bremen 39th place (471) out of the 44 countries and states (Prenzel et al. 2005, p. 60). Similarly, in terms of the slope of the socio-economic gradients that describe the extent to which students' test scores depend on their socio-economic background, the state of Brandenburg is the 10th most equal unit (slope 37.5), while the state of Mecklenburg-Western Pomerania takes 40th place (48.3) (Prenzel et al. 2005, p. 254; OECD 2004, p. 397).

The state-level nature of German education policy also provides for substantive institutional variation between German states. This is particularly true for central exit exams and tracking. In 2003, seven of the 16 states had external exit exams at the end of secondary school, while the other nine states had only exams devised, conducted, and graded by individual teachers. While most states tracked their students into different-ability schools after fourth grade (age ten), there are two states that conduct the tracking two grades later (age twelve). This does not bring them close to the OECD median of first tracking at age fifteen (cf. OECD 2004, p. 262), but it might still have an effect on student outcomes. Furthermore, the number of school types and thus the extent of tracking range from two to four types. There also is noteworthy variation in expenditure per student, the share of private schools, and facilities for early childhood education. By contrast, German states are rather uniform in allowing relatively little autonomy for their schools.

[4] At the time of writing, student-level data of the PISA-E 2003 test have not yet been released. Some student-level data of the PISA-E 2000 test have been released, but due to apparent confidentiality requirements, this dataset does not include state identifiers, school identifiers, and most of the relevant background variables. As a result, there is no way to perform analyses of German cross-state variation in educational outcomes using student-level data.

[5] Of course, the use of student-level data would not change the fact that the number of observations on school policies is only 16; it would only allow for more extensive controlling for individual student and school factors.

Economic theory suggests that these kinds of institutional frameworks of school systems may be fundamental determinants of the efficiency and equity of learning outcomes.[6] Because institutions frame the incentives that are at play in the school system – the explicit and implicit rewards for teaching and learning, as well as the consequences of poor teaching and learning – they will determine the behavior of students, parents, teachers, schools, and administrators.

The public discussion on the merits of the different institutional features is very intensive in Germany. However, it tends to be rather ideological, and if data are referred to at all, it tends to be in the form of univariate comparisons of two select states. For example, casual panelists would do away the cross-country evidence on detrimental effects of early tracking on equality of opportunity by stating that Bavaria (the most equal early tracker) achieves more equality than Berlin (one of the two states with delayed tracking). No-one seems to have analyzed the associations between the different institutional factors and the substantial variation in efficiency and equity of school outcomes across the 16 states in a multivariate and systematic way.

Given this dearth of analysis, the results reported in this paper constitute a significant advancement of our understanding of the associations between institutions and outcomes in the school system. At the same time, they should be interpreted cautiously as controlled descriptive associations. Even if biases from unobserved country heterogeneity can be ruled out in the within-German variation, an interpretation as causal effects might still be hampered by other forms of endogeneity, and the limited degrees of freedom preclude deeper analyses.

The main finding of the paper is that the German state-level analysis strongly confirms the cross-country evidence on the importance of institutional conditions for educational outcomes. External exit exams are systematically associated with higher student performance in math, science, and reading.[7] Delayed tracking and fewer school tracks are significantly related to higher equality of opportunity, in the sense that students' performance depends less strongly on their socio-economic background. The enrollment share in privately operated schools is associated with higher reading performance, as well as with higher equality of opportunity. While the socio-economic background is also strongly associated with the efficiency and equity of educational outcomes, expenditure per student and class size are not. As in the cross-country evidence, there are no signs of efficiency-equity tradeoffs. When we pool the German states with OECD

[6] For theoretical models of the effects of different institutions, see Bishop and Wößmann (2004) and Bishop (2006) for effects on educational efficiency and section 3 of Schütz et al. (2005) for effects on educational equity.

[7] In this paper, results are reported only for math; results for science and reading are available from the author.

countries in a joint analysis to test for consistency between German and international variation, all institutional effects are also significant across OECD countries, with the exception of private school operation in the equity model. Even more, the size of the effects is statistically indistinguishable between the German state sample and the OECD country sample. The fact that the German results are strongly in line with the international evidence lends support to the conclusion that the existing cross-country evidence may not be substantially biased after all.

The remainder of the paper is structured as follows. Section 2 presents the data. Section 3 discusses the empirical models. Sections 4 and 5 present the results on the efficiency and equity of educational production across German states, respectively. Section 6 reports results of models pooling German states with OECD countries. Section 7 concludes.

2. Schooling Data across German States

German state data on student performance are provided by a national extension (PISA-E) of the Programme for International Student Assessment (PISA), conducted by the Organisation for Economic Co-operation and Development in 2003 (cf. OECD 2004 for details). PISA tested representative samples of 15-year-old students in math, science, and reading literacy, with a focus of test items on real-life applications. PISA-E used the same tests as the international PISA, but additional testing ensured representative sampling within each of the 16 German states. PISA-E tested a total of 44,580 students in 1,487 schools, with state samples ranging from 1,618 students in Saxony-Anhalt to 4,904 students in Hamburg (cf. Prenzel et al. 2005 for details).

German states' mean performance is reported in the official PISA-E publication (Prenzel et al. 2005), measured on a scale with mean 500 and standard deviation 100 across all OECD countries. Because PISA 2003 had a special focus on math (taking over half of the assessment time), this paper focuses on math achievement.[8] On average, the math performance of German students is close to the OECD average, at mean 503 and standard deviation 103.

Prenzel et al. (2005) also report slopes of socio-economic gradients for each state. This is the first time ever that such information has been released for German states. The gradients are obtained by regressing students' math performance on their socio-economic background (cf. chapter 4 of OECD 2004). Socio-economic background is measured by the PISA index of economic, social, and cultural status (ESCS), which uses principal component analysis to

[8] In terms of effects on mean performance, the main qualitative results presented for math in this paper hold in the same way for science and reading. Socio-economic gradients are not available for science and reading.

combine measures of occupational status, education, and wealth (proxied by availability of relevant household items). Thus, measuring how much difference one unit on the ESCS scale (which is standardized to have an OECD mean of zero and a standard deviation of one) makes to student performance in math, the slopes of the socio-economic gradients can be interpreted as proxies for equality of opportunity for children from different family backgrounds.

Performance levels and slopes of the socio-economic gradients are reported in Table 1. It is evident that there is substantial cross-state variation in both dimensions. Figure 1 plots the two measures against each other, and Figure A1 in the appendix provides a map of the German states with a rough depiction of the two measures.

We combine the PISA data with data on socio-economic background, educational resources, and different institutional features of the school systems from several sources (see Table A1 in the appendix for sources, definitions, and descriptive statistics). Table 1 reports selected variables for the 16 German states, revealing substantial variation in several school policy measures.

3. Empirical Models

The empirical models are straightforward, taken directly from the literature on education production functions. For the efficiency estimations, mean test-score performance *T* of students in state *s* is regressed on the institutional features *I* of interest and a set of control variables *C*:

$$T_s = \beta_0 + I_s\beta_1 + C_s\beta_2 + \varepsilon_s \tag{1}$$

In the education production function literature, it is common that the control vector include at least one measure each of family background and school resources (cf. Hanushek 2002).[9] Given the control for material inputs, estimated effects of institutions on student outcomes can be interpreted in terms of efficiency in such a model.

For the equity estimations, the slope of the socio-economic gradient *S*, as described above, is similarly regressed on institutional features *I* and control variables *C*:

$$S_s = \alpha_0 + I_s\alpha_1 + C_s\alpha_2 + \eta_s \tag{2}$$

where both vectors may contain different sets of variables as above. Using cross-country data, Schütz et al. (2005) show that this specification yields the same qualitative results as a

[9] Cf. section III in Wößmann (2003a) and section 3.2 in Fuchs and Wößmann (2007) for extensive discussion of specification issues of such cross-sectional models.

specification using student-level data where the effects of institutions on the slope of the socio-economic gradient are identified by the coefficient on an interaction term between the system-level institutional features and the student-level measure of socio-economic background that is used in the estimation of the socio-economic gradient.

4. Efficiency of Educational Production

This section reports results of regressing mean student performance on different explanatory variables across German states. Results of estimating different specifications of equation (1) are presented in Table 2. In all specifications, the dependent variable is states' mean performance on the PISA-E 2003 math test, and the sample includes all 16 German states. We start by discussing the effects of external exit exams, students' socio-economic background, schools' resource endowments, and the existence of comprehensive schools, respectively. We then present evidence on the extent to which the estimated effects of external exit exams may reflect genuine effects of the institutional structure or rather effects of differences in political orientations. We close by discussing evidence for additional factors, including private schools and the pre-school system, as well as evidence on whether there is an efficiency-equity tradeoff.

The first column of Table 2 reports our basic model. When dealing with a sample of 16 observations, parsimoniousness of the model is a virtue. Therefore, we aim to keep our models as small as possible, while at the same time not leaving out any variables that show independent significant effects. As a consequence, after some specification checks with different variables of socio-economic background and educational resources, the basic model includes two measures of family background – an index of socio-economic background (ESCS) and the employment status of students' fathers – as well as a dummy on the existence of external exit exams. It turns out that once these three variables are included, no other available measure of family background and no available measure of school resources enter significantly in the model, and no additional variable changes the main findings on the included variables. Together, the three variables account for 83 percent of the total cross-state variation in student performance.

(1) External exit exams. At the time of testing, seven of the 16 German states had examination systems at the end of secondary school which require students to take external exams that are the same for all students in the state. By contrast, the other nine states had exit exams that were devised, conducted, and graded by individual teachers only. Economic theory suggests that external exit exams can function as accountability devices that change the behavior of students, parents, teachers, and principals in ways which improve students' learning outcomes (cf. Bishop and Wößmann 2004; Bishop 2006).

In line with the theory, external exit exams are statistically significantly related to better student performance across German states. The size of the estimated effect is substantial, at 17.9 PISA test-score points or 17.9 percent of an international standard deviation. This compares, for example, with an average increase in learning over one year of age of 28.5 test-score points in the German PISA data.[10] As another example, the effect size can account for 38 percent of the total difference of 46.6 points between the best-performing and the worst-performing non-city state, Bavaria (which has external exit exams) and North Rhine-Westphalia (which does not).

Figure 2 depicts the association graphically as an added-variable plot.[11] The states to the right, with a conditional value exceeding 0, are the ones that have external exit exams. Note that (conditional on socio-economic background) only one state without central exams, Berlin, outperforms at least one state with central exams (Baden-Wurttemberg and Saarland). With this exception, *any* state with central exams performs better than *any* state without central exams.

The positive association between external exit exams and student outcomes is very robust and survives all the robustness specifications that we report below.[12]

(2) Socio-economic background. The basic model includes two measures of family background, namely the share of students whose father has full-time employment and the PISA index of economic, social, and cultural status (ESCS), which combines measures of parental occupation, education, and wealth. Both measures enter statistically significantly positive in predicting student performance. The finding of substantial family-background effects is in line with the rest of the literature on education production functions (e.g., Hanushek 2002; Schütz et al. 2005; Wößmann 2003a). Figure 3 depicts the associations of the two family-background measures with student performance graphically. Once the employment status of the students'

[10] This is based on the coefficient of a bivariate regression of math test score on student age (which ranges from 15 years 2 months to 16 years 4 months) using German student-level data from the international PISA sample.

[11] Added-variable plots show the association between two variables after the influence of control variables is taken out. Thus, both variables are first regressed on the controls. Only the residuals of these two regressions, which reflect the parts of the variation in the two variables that cannot be accounted for by the controls, are used in the graph. This is numerically equivalent to including the controls in a multivariate regression, as in Tables 2 and 3.

[12] Cf. Jürges et al. (2005a; 2005b) for similar evidence of positive effects of external exit exams across German states, applying differences-in-differences and matching techniques to data on 14 German states from the 1995 Third International Mathematics and Science Study (TIMSS) and the 2000 PISA-E study.

fathers is included, the average unemployment rate of the state does not enter significantly and leaves the estimate on fathers' employment significant.[13]

Columns (2)-(4) of Table 2 report specifications that add several additional background measures to our basic model. First, we add two dummies for city states and for East German states. After controlling for the variables of the basic model, both dummies are positive, but statistically insignificant (individually, jointly, and when added individually). The next specification adds a states' gross domestic product (GDP) per capita and parental education. Again, when added to the basic model, both additional measures of socio-economic background do not enter significantly (individually, jointly, or when added individually).

In specification (4), the share of students with migration background, defined as having at least one parent born in a foreign country, does not enter the model significantly. The same is true with alternative definitions of migration background such as immigrant families (student and both parents born abroad), first-generation immigration (both parents born abroad, but students themselves not), or only one parent born abroad. Different measures of family structure, such as the share of students living with a single parent, in a blended family, or with both biological parents, also do not enter the model significantly.

(3) Resources. Columns (5) and (6) of Table 2 add measures of schools' endowments with resources to the basic model. The first specification uses expenditure per student in the state as an encompassing financial aggregate of the educational resources available. It is not statistically significant, and has a negative point estimate.[14] The next specification uses two measures of real classroom resources, namely a state's average class size and average instruction time in math.[15] Both measures are not statistically significant (individually, jointly, or when entered individually), and the point estimate on class size is counterintuitively positive. Figure 4 plots the conditional associations of student performance with the two most conventional resource measures, expenditure per student and class size. In both cases, it is clear that the lack of a significant association is not driven by outliers or obvious non-linearities.

[13] Similarly, once the ESCS index is included, an alternative index of socio-economic background provided by PISA, the International Socio-Economic Index (ISEI), which is based on the status of parents' occupations and provides part of the information entering the construction of the ESCS index, does not enter significantly and leaves the ESCS index significant.

[14] The distribution of educational expenditure across states is very stable over time. Using a lagged rather than contemporaneous value of expenditure, e.g. for 2000, yields the same result.

[15] Note that teacher education levels do not vary across states in terms of length of study, with all states requiring a university teacher education to become a teacher.

Note that when measuring resources at the state level, biases from non-random placements of students between and within schools, which hamper conventional micro estimates (cf. Wößmann and West 2006), should no longer be relevant. At the same time, Fertig and Wright (2005) use the international PISA 2000 data to corroborate the US evidence by Hanushek et al. (1996) that the probability of finding significant correctly-signed class-size effects increases with the level of aggregation of the class-size measure. Despite this likely aggregation bias, we do not find significant effects of class size on student performance across German states.[16]

(4) Comprehensive schools: efficiency or selection? An obvious institutional feature that distinguishes German states from one another is the existence of comprehensive schools (Gesamtschulen). Half the states have comprehensive schools, the other half does not. The specific aspect of German comprehensive schools is that they do not exist *instead of* a selective system of different schools tracked by ability, but that they are rather added *next to* the existing tracks (usually Gymnasium, Realschule, and Hauptschule) in the states that introduced them. The PISA-E data have revealed that at the school level, comprehensive schools tend to perform much worse than expected, usually between the second and third track schools (Prenzel et al. 2005). However, from the descriptive pattern it is not clear whether the comprehensive schools indeed do a bad job in teaching their students, or whether it is the case that relatively poorly performing students tend to select themselves into the comprehensive schools. In particular, there is a feeling that highly ambitious parents send children whose primary-school performance suggests that they will not succeed in the highest-track schools to comprehensive schools, rather than to lower-track schools. That is, selection is likely based on student characteristics that usually go unobserved. While such selection would not affect states' mean performance, a poor performance of comprehensive schools in teaching skills should drag down state mean performance.

Casual observation and general public discussion seem to suggest that the latter is true: As specification (7) of Table 2 shows, states with comprehensive schools perform significantly worse, even after controlling for family background. However, to a noteworthy extent, the states with comprehensive schools also tend to be the states that do not have external exit exams: the correlation coefficient is significantly negative at -0.63. A "horse raise" specification that includes both comprehensive schools and external exit exams provides a clear picture which of

---

[16] The lack of significant effects of resources in Germany is corroborated by time-series evidence (Gundlach et al. 2001) and quasi-experimental micro evidence (Wößmann 2005a).

the two is relevant for a state's mean performance. As specification (8) shows, the significantly positive effect of external exit exams is hardly changed when the comprehensive school dummy is introduced, while the negative coefficient on the comprehensive school dummy gets much smaller (in absolute terms) and statistically insignificant once external exit exams are in the model. It seems that holding other relevant influences constant, states that added comprehensive schools to their existing tracked system perform only slightly, and not statistically significantly, worse than states that did not introduce the comprehensive school type.

Similar to the comprehensive school dummy, other measures of tracking such as the age at which the first tracking occurs and the number of tracks do not enter significantly into the model. However, both measures have consistently negative coefficients, suggesting that if anything, the extent of tracking is negatively associated with a state's mean student performance.

(5) Institutions or values? In Germany, there is a clear tendency for politically conservative states both to have external exit exams and not to have comprehensive schools. This raises the question whether it is indeed the institutional feature of external exit exams which is related to student outcomes, or whether it is rather conservative values or governments. We obtained two measures of conservativeness, one referring to the specific state government, the other to the populace. The first measure is a dummy whether the respective state prime minister, and thus the main ruling party in the state, belongs to the conservative party. In 2003, this was true for nine of the 16 states. The correlation of conservative state government with external exit exams is 0.52, and -0.38 with comprehensive schools. The second measure is the share of conservative votes in the German federal election of September 2002. Referring to the same point in time and the same federal government to vote (rather than specific state governments), this measure provides a proxy for the political values of the population in general. The correlation of the conservative vote with external exit exams is 0.47, and -0.54 with comprehensive schools.

Specification (9) of Table 2 adds both measures of conservativeness to our basic model. The first result is that there remains a statistically significant effect of external exit exams, only slightly reduced in magnitude. That is, there seems to be an effect of the institutional feature of external exit exams, independent of the political orientation of the government or the populace. The second result is that there is a significant positive association of student performance with the share of conservative votes in the federal election, but not with whether the respective state government is a conservative one or not (with or without controlling for the share of conservative votes). By adding the share of conservative votes, the cross-state variance in student performance explained by the model increases to 91 percent (the adjusted R^2 to 0.86). While the political variables are entered as control variables and not necessarily for causal interpretation, it seems that there is something about a conservative population, rather than about a conservative government, which is conducive to student performance. For example, parents with conservative values may have a higher valuation of high-quality education in general. However, mostly independent of that, external exit exams exert a significant effect on student performance.[17]

(6) Private schools, pre-school system, and other additional factors. At 6 percent on average, the share of students attending privately managed schools in Germany is relatively small compared to the OECD mean. Still, the private school share varies across states from 2.2 percent in Brandenburg to 12.2 percent in Bavaria. Economic theory suggests that private school operation may increase the efficiency of a school system because market forces create incentives for cost effectiveness and performance-conducive qualitative innovation, both for privately managed schools and for nearby public schools that face their competition (cf., e.g., Shleifer 1998; Bishop and Wößmann 2004). International evidence suggests that the share of privately operated schools is indeed positively related to a country's student outcomes (Wößmann 2005b; cf. also Toma 2005). However, Weiß and Preuschoff (2006) have argued that such an effect may be less relevant in Germany, because the majority of privately operated schools are subject to tight public regulation which makes most of their conduct closely resemble public schools (cf. Toma 1996).[18] As specification (10) of Table 2 shows, the private school share is not statistically significantly related to a state's mean math performance, although the point estimate is positive (same when entered without pre-school enrollment). But this is the only occasion where results differ across subjects: There is a statistically significant positive association between the share of students attending private schools and a state's mean reading performance (coefficient 1.42, standard error 0.60) when the private school share is added to the basic model (cf. Figure 5).

Cross-country evidence suggests that the duration of the pre-school cycle (albeit not necessarily pre-school enrollment) may be positively associated with student performance (Schütz et al. 2005). While the official duration of pre-school cycles does not vary across German states, pre-school enrollment does. The share of children aged 3 to 5 who are enrolled in day-care centers or other pre-school facilities is not significantly associated with student performance in

[17] The comprehensive school dummy remains insignificant in the specification with the political measures.

[18] Unfortunately, there is no data available on differences in the public funding for private schools across German states, which proved to be important in cross-country analyses (Wößmann 2003a, 2005b).

our model (specification (10) of Table 2; same when entered without private school enrollment).[19] However, pre-school enrollment is strongly correlated with external exit exams across German states (correlation of 0.69). As long as external exit exams are not included in the model, pre-school enrollment is significantly positively associated with student performance.

Another institutional feature that is significantly related to student outcomes in cross-country research is the extent of autonomy of schools in different decision-making areas (e.g., Wößmann 2003a, 2003b). But German states are rather uniform in allowing relatively little autonomy for their schools, in particular in personnel matters. To the extent that there may be some differences, there is no comparable data available.

The German PISA-E report provides state-level data on hours of math homework, on the share of students receiving private afternoon lessons and receiving other supplementary lessons, and on the share of students coming late to class. It also reports the share of students in a state whose enrollment in school was delayed by one year and who repeated a grade in primary or in secondary school. None of these measures enters significantly into our basic model.

Finally, we add controls for religious affiliation to our basic model. Across the German states, religious affiliation may be historically associated with cultural differences and valuation of education. The share of Protestants, Catholics, and persons not affiliated with these two main religious associations in the total population, as well as the share of Protestants in the Protestant-plus-Catholic population, all do not enter significantly. The same is true when using historical values on religious affiliations from the 1950s to 1980s, available for the West German states.

(7) Efficiency-equity tradeoff? A lot of public discussion surrounds the worry that there may be a tradeoff between achieving efficiency and equity in educational outcomes. The descriptive pattern of Figure 1, however, already suggested that this hardly seems to be the case. As specification (11) of Table 2 shows, the slope of the socio-economic gradient, and thus the extent to which equity is achieved in the sense that student performance does not depend on family background, is not significantly related to mean performance. The point estimate is negative, suggesting that higher inequality is related to lower mean performance. Similarly, institutional measures which will be shown below to be associated with the slope of the socio-economic

[19] The available data on pre-school enrollment refer to school year 2005/06. This is the first time that data on actual pre-school enrollment were collected. The only data previously available (for 2002 for the last time) are on the number of officially available places in day-care centers and nurseries (Statistisches Bundesamt 2004), and it is well known that there are huge discrepancies between states' official numbers of slots and actually filled places.

gradient, namely age of first tracking and number of tracks, are unrelated to mean performance. Thus, the performance level does not have to be traded off against equity.

5. Equity of Educational Production

This section reports results of regressing the slope of socio-economic gradient on different explanatory variables across German states. Results of estimating different specifications of equation (2) are presented in Table 3. In all specifications, the dependent variable is the slope of the socio-economic gradient in the PISA-E 2003 math test of each state, which measures to what extent individual students' performance depends on their socio-economic background and thus serves as a measure of inequality of opportunity for students from different family backgrounds. We start by discussing the effects of delayed tracking and the number of tracks, including the existence of comprehensive schools and the role of differences in political orientations. We then report results for several measures of students' socio-economic background and additional factors, including schools' resource endowments, the pre-school system, and private schools. We close again with evidence on whether there is an efficiency-equity tradeoff.

The first column of Table 3 reports our basic equity model. It includes two measures of family background – GDP per capita and the standard deviation of the ESCS index of socio-economic background – as well as two measures of the extent of tracking – the number of school tracks in the state and a dummy for states that track their students later than is standard. It turns out that once these four variables are included, no other available measure of family background and no available measure of school resources enter significantly in the model, and no additional variable changes the main findings on these variables (with the possible exception of the number of tracks, which sometimes drops in significance). Together, the four variables account for 67 percent of the total cross-state variation in our measure of inequality of opportunity.

***(1) Delayed tracking*.** In Germany, it is standard that students are tracked, based on their prior performance, into different types of school after grade four (at age ten). There are two states, though – Berlin and Brandenburg – where the tracking takes place two years later, after grade six. That is, in these states all students are taught together for two more years in primary school.[20] While the predictions of economic theory on the effects of tracking on student outcomes

---

[20] There are four additional states which have a so-called "orientation phase" in grades five and six where students can still decide which track is best for them. We did not find any significant difference in the efficiency or equity of student outcomes between these states and the states with pure early tracking.

heavily depend on the type of peer effects assumed, there is a tendency for models to find that later tracking is beneficial to equality of opportunity (cf., e.g., section 3 of Schütz et al. 2005). International evidence suggests that later tracking is indeed associated with increased equality in educational outcomes (Hanushek and Wößmann 2006; Schütz et al. 2005).

We find a similar effect across German states. The slope of the socio-economic gradient is significantly lower in states with later tracking. This effect is highly robust to the multitude of robustness specifications reported below. The size of the effect, with a slope of the socio-economic gradient that is 5.1 units lower in late-tracking states, is 1.6 times as large as a standard deviation in the slope measure, and accounts for nearly half the slope difference of 10.8 between the most unequal (Mecklenburg-Western Pomerania) and the most equal state (Brandenburg).

As the added-variable plot of Figure 6 shows, holding the control variables constant, the two states with delayed tracking are both more equal than any of the other 14 German states. Of course, the effect is identified from two states only. But at least, it is worth noting that the results hold when we drop either of the two states from the sample. That is, both states strongly support the positive effect of later tracking on equality of opportunity.[21]

From Figures 5-7, which depict the associations of the basic model, it is apparent that Rhineland-Palatinate (RPF) is a noteworthy outlier to the bottom in all four dimensions. As specification (13) of Table 3 shows, the reported results get substantially stronger once this state is dropped from the sample. Note that the outlier quality of Rhineland-Palatinate is confirmed in robust regression techniques that drop or downweight outliers based on statistical indicators.[22] The applied robust regression model attributes by far the lowest weight to Rhineland-Palatinate. Our reported results are robust to using the robust regression estimator.[23] In the remaining specifications, we keep including Rhineland-Palatinate in the sample, but it should be noted that all reported results are perfectly robust to dropping it from the sample.

(2) Number of tracks and comprehensive schools. In addition to the age of first tracking, there also is variation across German states in the number of tracks, or different school types.

---

[21] Bauer and Riphahn (2006) present evidence from across Swiss cantons which similarly shows that later tracking increases educational mobility.

[22] The specific method used, corresponding to the *rreg* robust estimation command in *Stata*, eliminates gross outliers for which Cook's distance measure is greater than one (which is not the case for any observation in our applications), and then iteratively downweights observations with large absolute residuals.

[23] Detailed results are available from the author. Note that the efficiency specifications of Table 2 are also perfectly robust to using robust regression techniques. In this case, Bavaria and Saxony get downweighted. The results reported there are also robust to dropping these two states from the analyses altogether.

Four states have only two selective tracks, while the standard classical model (at least in West Germany) is to have three tracks. In addition to that, as described above, several states have added comprehensive schools as a quasi fourth type next to the other tracks, in effect yielding a total of four tracks. By increasing selectivity, the number of tracks may be viewed as an additional aspect of the intensity of tracking.[24] International evidence suggests that a larger number of school types increases inequality of opportunity (Ammermüller 2005).

In our analysis, we measure the number of tracks in two variables, a dummy for the existence of comprehensive schools (Gesamtschulen) and a variable on the number of the remaining school types. The division is undertaken in order to distinguish between effects of comprehensive schools, which constitute the peculiar case of a non-tracked additional school type, and the number of tracks in general. However, results on the number of tracks are very similar if we count comprehensive schools as an additional track in the number-of-tracks variable.[25]

As specification (14) of Table 3 shows, the number of tracks is significantly positively associated with the measure of inequality of opportunity, while the existence of comprehensive schools is not significantly associated with it. Each additional non-comprehensive track increases the slope of the socio-economic gradient by 2 units, or 64 percent of a standard deviation of the slope measure. Given the insignificance of the comprehensive school dummy, we have dropped it from our basic model specification (12), as well as from the following robustness specification. It should be noted that everything that follows is perfectly robust to including the comprehensive school dummy, which stays statistically insignificant in all the models.[26]

Figure 7 depicts the association between the number of tracks and inequality of opportunity graphically. Two states, Mecklenburg-Western Pomerania (MVO) and Rhineland-Palatinate (RPF), appear as outliers to the right of this graph. When both are dropped from the sample, the

---

[24] Using student-level data from the 2000 PISA-E study, as well as longitudinal data from two German states, Baumert et al. (2006) show that the type of school track is indeed associated with differential learning, even after controlling for a measure of basic intellectual competencies (or earlier test scores in the longitudinal data) and compositional effects (cf. Köller and Baumert 2001).

[25] Of course, as long as the comprehensive school dummy is included, coefficient estimates on the number-of-tracks variable with and without counting comprehensive schools are numerically equivalent. The only difference between the two models is in the estimate of the comprehensive school dummy, which turns out to be statistically insignificant in all our models anyways.

[26] It should also be noted that a consistent definition of number of tracks is not straightforward across states. We have decided to use official statistics on the number of different school types which have at least five percent of the student population each. Other possible definitions, e.g. based on narrative accounts from the official PISA-E publications, also produce positive coefficient estimates, which are not always statistically significant, though.

results of the basic model only get stronger, with the coefficient estimate on the number of tracks increasing to 2.92 (standard error 1.45).

To an extent, the number of tracks is an indicator of the extent to which students in lower-track schools are excluded. If there are many tracks, the lowest-track school is likely to cater only a relatively small fraction of students, which are then separated from the core of the population, at the bottom. Thus, an alternative measure of the intensity of tracking would be the share of students attending the lowest existing track in a state. It turns out that this measure yields the same pattern of results when used instead of the number of tracks – the share of students in the lowest track is significantly associated with lower inequality of opportunity. When both are entered jointly, neither dominates the other, and they are jointly (but not individually) significant. Thus, the statistical analysis cannot distinguish between these two alternative measures, and we interpret them as proxying for the same underlying concept, the intensity of tracking and the exclusion of low-performing students that is connected with it.[27]

Specification (15) adds our two variables of political orientation. Neither the indicator of a conservative state government nor the share of conservative votes in the federal election enters significantly in the model (individually, jointly, or when entered individually). It seems that the effects of tracking are again true effects of the institutional structure of the school system, rather than proxies for the political orientation of the government or the populace.

Finally, states differ in terms of whether the decision about which track to enter is firmly decided based on marks in primary school or whether parents have a lot of discretion in the decision. An indicator for states where marks decide on the transfer enters the model consistently negatively, but never statistically significantly.

***(3) Socio-economic background*.** Two measures of socio-economic background are included significantly in the basic model. The first one is the standard deviation of the PISA index of economic, social, and cultural status (ESCS), which was used in the construction of the socio-economic gradients. A higher inequality in the socio-economic background of the parents is associated with a steeper slope of the socio-economic gradient.[28] The second measure is a state's GDP per capita, which is negatively associated with the slope of the socio-economic gradient.

[27] A third alternative measure, a dummy for whether the Hauptschule track exists in a state, has similar properties to the other two measures.

[28] Very similar results are obtained when the difference between the 95th and the 5th percentile in the ESCS is used instead of the standard deviation in the ESCS; the correlation coefficient between the two is 0.989.

That is, more economically advanced states tend to show lower inequality of opportunity. The two associations are depicted graphically in Figure 8.

Specification (16) of Table 3 adds indicators for city states and for states in East Germany. Both are again statistically insignificant once the variables of the basic model are included. The same is true for the following alternative measures of socio-economic background: the ESCS index, the ISEI index, fathers' employment status, the state's unemployment rate, and indicators of family structure. The education level of parents tends to enter significantly positive in some models, but does not change any of the reported results. Different measures of migration background enter negatively, but statistically insignificantly, although some get close to the margin of standard significance levels (sometimes reaching a 15 percent level).

(4) Resources, pre-school system, private schools, and other additional factors. A state's average educational expenditure per student and instruction time do not enter significantly in the basic equity model.[29] Class size enters significantly negatively in the basic model (specification (17) of Table 3) – smaller class sizes being associated with larger inequality – but this association loses statistical significance once the only other significant addition to the model, the share of private schools (see below), is included.[30]

Pre-school enrollment of children aged 3 to 5 is negatively associated with the slope of the socio-economic gradient (specification (18)), but not statistically significantly at conventional levels (albeit at a 25% level).[31] The coefficient approaches statistical significance (at the 15% level) when only children aged 4 are considered. While not conclusive, such an effect would be in line with the international evidence of an equality-enhancing effect of extensive pre-school systems (Schütz et al. 2005). The somewhat weaker association in Germany might be due to the fact that German pre-school institutions tend to be viewed as mere day-care places which overlook the children, without explicit educational mandates.

The share of students whose enrollment in school was delayed by one year is positively associated with inequality of opportunity (specification (19)), which may suggest that the practice

---

[29] Unfortunately, no indicators are available on the variation in resources available to students with high and low socio-economic background.

[30] Private schools tend to be less well funded than public schools, giving rise to a positive correlation between class size and share of private schools.

[31] There is substantial correlation (-0.46) between pre-school enrollment and the standard deviation of socio-economic background, and the negative coefficient on pre-school enrollment gets significant in a model that excludes the latter variable.

of late enrollment tends to hurt particularly children from low socio-economic backgrounds. However, the association is not statistically significant, although it gets close to statistical significance in some models, for example in the sample without RPF.

The only additional factor that consistently enters the equity model significantly is the share of students attending privately operated schools. As reported in specification (20) of Table 3, the private school share is significantly negatively associated with the slope of the socio-economic gradient (cf. Figure 9). That is, a larger private school share is associated with more equality of opportunity. This finding mirrors similar international evidence, which at the same time controls for the share of private funding (Schütz et al. 2005). While the latter information is not consistently available across German states, it should be noted that most privately managed schools in Germany do not require private fees, but receive most of their funding from the state.[32] A possible explanation for the finding in the German setting is that parents with low socio-economic background may be more likely to send their children to higher-track schools if the schools are religious schools, which may cater better to their traditional values. In fact, most of the privately managed schools in Germany are run by the (Catholic or Protestant) church, and a majority of them are schools of the highest, academic track. These schools may make it more acceptable for parents who have low education themselves to send their children to high-track schools, where they tend to learn higher standards, while these parents may be averse to standard high-track schools because of their differing aspirations and because of a feeling that they cannot support their children at those schools. Note that including the private-school measure increases the variation in inequality of opportunity explained by the model to 84 percent (the adjusted R^2 to 0.76; in the sample without RPF to 0.92 and 0.87, respectively).

The fraction of students who repeated a grade is not significantly associated with the slope of the socio-economic gradient. The same is true for the share of students receiving private or other supplementary lessons. As currently implemented in Germany, these features do not seem to have first-order effects on the equality of opportunity achieved. Finally, as in the efficiency model, the controls for religious affiliation do not enter significantly in the basic equity model.

(5) Efficiency-equity tradeoff? The main institutional feature significantly associated with mean performance, external exit exams, is not significantly related to equity (specification (21)).

[32] The average share of public funding in privately operated schools in Germany is 77.1 percent in PISA 2000 (Wößmann 2005b), and the remainder usually comes from church sources, but not from parental fees; unfortunately, this information is not available at the state level.

Likewise, mean performance itself does not enter the equity model significantly (specification (22)). Again, there is no evidence of any tradeoff between efficiency and equity.

6. Pooling German States with OECD Countries

The basic patterns of the presented evidence from cross-state variation within Germany confirm the available cross-country evidence. For a more thorough comparison, we can test the consistency of the German and the international evidence within the same framework and data by performing the same models for the aggregate countries participating in PISA 2003. Such an approach does not only allow for a test whether the German results can be replicated with international variation, but also for statistical tests of whether results differ significantly between German and international variation. Table A2 in the appendix reports descriptive statistics for the sample of OECD countries.

In terms of population size, it is not inappropriate to combine German states with other OECD countries. For example, the three biggest German states (North Rhine-Westphalia at 18 million, Bavaria 12.5, and Baden-Wurttemberg 10.7) are each bigger than the median OECD country, or than such countries as Denmark (5.4), Switzerland (7.5), Austria (8.3), Sweden (9.1), and Belgium (10.5). Five of the 16 German states are bigger than the top performing OECD country in PISA, Finland (5.3). And even the smallest German state, the city state of Bremen at 660,000, is bigger than Iceland and Luxembourg (below 500,000).

To estimate the within-German and cross-country variation within the very same framework, we use the basic model specifications obtained above. First, we estimate the basic efficiency and equity models in the sample of 28 OECD countries (excluding Germany),[33] following studies such as Bishop (1997, 2006) and Lee and Barro (2001) in performing the analysis at the country level.[34] Next, we pool the German state and OECD country data to test for differences across the state and country sample. Finally, we add non-OECD countries to the model.

(1) Efficiency**.** Results of running the basic efficiency model on the non-German OECD sample are reported in specification (23) of Table 4. Comparing this to specification (1) of Table 2, it is obvious that the basic patterns are the same as in the sample of German states. In particular, external exit exams are again significantly positively related to student performance. It

---

[33] Due to a low response rate, PISA 2003 data have not been reported for the United Kingdom.

[34] Ongoing research (Wößmann et al., in process) will use the international PISA 2003 data at the individual student level, but German confidentiality practices prevent an identification of states in the student-level data.

may be noted that the precision of the estimate is lower in the OECD sample than in the German sample, suggesting that there is less noise among the more homogenous German states.

To test whether there are any significant differences between the German state sample and the OECD country sample, specification (24) pools both samples and includes a dummy for Germany, as well as interaction terms between the German dummy and all explanatory variables. Neither the German dummy nor the interactions are significant, either individually or jointly (cf. the *F*-test reported at the bottom of the table). That is, the data generation process across OECD countries seems to be the same as the one across German states. While larger heterogeneity may yield less precise estimates across countries, it does not introduce any significant bias.

As a consequence, we can estimate the model in the pooled sample of 44 data points without the insignificant interaction terms (specification (25)). Results are qualitatively the same as in the German state sample: Student performance is significantly higher in systems with external exit exams, with a higher index of socio-economic background, and with a larger share of employed fathers. Specification (26) reveals that class size is again insignificant when added to the basic model (also jointly with an interaction term with the German dummy if included). The same is true when adding educational expenditure per student.

Across German states, the share of students attending privately operated schools was positively, but insignificantly related to student performance in math, although the association was significant in reading. When we add the private school share to the pooled model in specifications (27) and (28), the positive association is statistically significant also in math, and the size is statistically indistinguishable between the samples of German states and OECD countries. Thus, the lack of statistical significance in the German sample seems to stem from the limited statistical power due to the small number of observations.

While the results reported so far refer to the relatively homogenous sample of economically advanced OECD member states only, several non-OECD countries participated in PISA 2003. As specifications (29) and (30) reveal, our substantive results are robust to including the non-OECD countries with available data in the analysis, which increases the sample to up to 54 data points. Interaction terms with a German dummy are again insignificant.

As in the German state sample, the indicator of late tracking, the number of tracks, and the slope of the socio-economic gradient again enter insignificantly in the pooled sample.

(2) Equity. Results of running the basic equity model on the non-German OECD sample are reported in specification (31) of Table 5. Inequality of opportunity is again negatively associated with the indicator of delayed tracking, and positively with the number of tracks. Both estimates are individually statistically significant at the 15 percent level only, but jointly at the 5 percent level (see the *F*-test of joint significance of the group of coefficients indicated by [g], reported at the bottom of the table). When pooling the German and the OECD sample (specification (32)), the German dummy and the interactions are not jointly statistically significant. However, the German dummy and its interaction with the standard deviation of the index of socio-economic background are individually significant, while the other interactions are far from statistical significance individually and jointly. Thus, specification (33) includes only the former two, which are now jointly significant at the 10 percent level.

The main finding of this basic pooled specification is that equality of opportunity is higher in systems with later tracking and with fewer tracks, and that these effects are not significantly different between the German state sample and the OECD country sample. Furthermore, while inequality of opportunity is significantly lower the larger the dispersion of socio-economic background across OECD countries, the opposite is true across German states. Controlling for this significant interaction, the dummy for Germany is significantly negative (it is not in a model without the interaction term). Finally, in the pooled model GDP per capita is no longer significant. While the difference in associations with indicators of socio-economic background is noteworthy, they are included and interpreted as control variables only in our models.

An indicator for German-type Gesamtschulen (comprehensive schools) remains statistically insignificant in the pooled regression. The qualitative result on the number of tracks is not affected by whether Gesamtschulen are counted as a separate track in the number of tracks or not.

The variation in the delay of tracking across OECD countries goes beyond the dichotomous German variation of whether students are tracked before age 12 or not. When using the age of first tracking as an alternative, continuous measure of the time of tracking, results are similar. However, when the delayed-tracking dummy is included in a "horse race" specification together with the continuous age-of-tracking variable, the dummy dominates.

As shown by specification (34), the negative association between private school enrollment and inequality of opportunity is specific to the German sample and does not show up in the OECD sample. This result is not affected by including the average share of public funding in the cross-country specification. The German-specific result may reflect the German-specific explanation suggested above that religious schools may attract students with low socio-economic background into high-track schools.

Specification (35) shows that neither mean performance nor external exit exams enter the equity model significantly, either individually or jointly. As in the German sample, the pooled sample does not reveal any signs of efficiency-equity tradeoff. Finally, specification (36) adds non-OECD countries with available data to the model. Unfortunately, only data on the age of first tracking, but not on the number of tracks are available for them. The result on delayed tracking is robust in the expanded sample, and in this sample it is even stronger when delayed tracking is measured continuously by the age of first tracking.

In sum, the main fundamental institutional determinants of interest in our models – external exit exams and private school shares in the efficiency model and delayed tracking and the number of tracks in the equity model – are similarly significant across OECD countries as they are across German states. Even more, their point estimates do not differ significantly between the two samples. The only exception of a significant difference is the share of private schools in the equity model, which is significant in the German model but not in the OECD model.

7. Conclusions

There is substantial variation in the level and inequality of student performance and in the institutional structures of the school systems across German states. The evidence provided in this paper shows that there are significant and robust associations between the institutional features and the efficiency and equity of educational outcomes. The existence of external exit exam systems is systematically related to higher student performance. Given that this result holds when controlling for resource inputs, it can be interpreted as an association between external exit exams and the efficiency of educational production. At least in reading, the share of privately operated schools is also positively associated with student performance.

Early tracking and the intensity of tracking (measured by the number of school tracks or the share of students attending the lowest-track school type) are both systematically related to a stronger dependence of students' performance on their socio-economic background. Thus, there is a significant and robust association between delayed, less intensive tracking and greater equality of opportunity. There also is a positive association between the share of privately operated schools and equality of opportunity in Germany. Although only at the verge of statistical significance, the same is true for larger pre-school enrollment. Also in line with the cross-country evidence, socio-economic background is strongly associated with student outcomes across German states, while the resource endowment of schools is not.

In terms of all three institutional determinants of school efficiency and equity – central exams, tracking, and private schools – the German evidence confirms existing cross-country evidence. When estimating the same models with the same PISA 2003 data on the sample of OECD countries, we find the same results (with the exception of private schools in the equity model): Countries with external exit exams and larger shares of privately operated schools have higher student performance, and countries with later tracking and fewer number of school tracks have higher equality of opportunity. What is more, the size of each of these institutional effects is statistically indistinguishable between the German state sample and the aggregate country sample in a pooled estimation of up to 54 observations. In a statistical sense, there is no difference in the fundamental determinants of the efficiency and equity of learning outcomes between the cross-state variation in Germany and the international variation across countries, providing additional validity to the results from across the 16 German states. In this sense, educational production in the German states provides a microcosm for the OECD countries.

Additional specifications of the German sample show that it is the institutional features of the school systems themselves, rather than the political orientation of the government or the populace with which some of them are correlated, which give rise to the association with student outcomes. There also is some association between political orientation and student outcomes independent of schooling institutions, but it is that mean performance tends to be higher in states with a larger population with conservative values, rather than with a conservative government.

The mere existence of comprehensive schools, which have been added next to the existing school tracks in some German states, is not significantly associated with either the efficiency or equity of student outcomes across the states.

Finally, there are no signs of efficiency-equity tradeoffs in educational production, either in the German or the international sample. There is no significant association between the performance level and the slope of the socio-economic gradient, and if anything, lower inequality is found in states with higher mean performance. There are also no cross effects of the main institutional features: central exams are associated with efficiency, but not with equity, while tracking is associated with equity, but not with efficiency. The only exception is private school operation, which is positively associated with both efficiency and equity (the latter at least in Germany). This again suggests that if anything, there is a complementarity rather than a tradeoff between achieving efficiency and equity in education.

References

Ammermüller, Andreas (2005). Educational Opportunities and the Role of Institutions. ZEW Discussion Paper 05-44. Mannheim: Centre for European Economic Research.

Bauer, Philipp, Regina T. Riphahn (2006). Timing of School Tracking as a Determinant of Intergenerational Transmission of Education. *Economics Letters* 91 (1): 90-97.

Baumert, Jürgen, Petra Stanat, Rainer Watermann (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In: Jürgen Baumert, Petra Stanat, Rainer Watermann (eds.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000*: 95-188. Wiesbaden: VS Verlag für Sozialwissenschaften.

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

Bishop, John H. (2006). Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn – Externalities, Information Problems and Peer Pressure. In: Eric A. Hanushek, Finis Welch (eds.), *Handbook of the Economics of Education*, Vol. 2: 909-944. Amsterdam: North-Holland.

Bishop, John H., Ludger Wößmann (2004). Institutional Effects in a Simple Model of Educational Production. *Education Economics* 12 (1): 17-38.

Bundeswahlleiter (2002). Wahl zum 15. Bundestag am 22. September 2002. Heft 3: Endgültige Ergebnisse nach Wahlkreisen. Wiesbaden: Statistisches Bundesamt.

Demmer, Marianne (2004). Kinder, Eltern, Lehrkräfte: Die frühe Schulformentscheidung in Deutschland überfordert alle Beteiligten. Frankfurt am Main: Gewerkschaft Erziehung und Wissenschaft. http://www.gew.de/Binaries/Binary6311/schulformentscheidung.pdf [13.2.2007].

Evangelische Kirche in Deutschland (2004). Kirchenmitgliederzahlen am 31.12.2003. http://www.ekd.de/download/kimi_2003.pdf [17.3.2007].

Fertig, Michael, Robert E. Wright (2005). School Quality, Educational Attainment and Aggregation Bias. *Economics Letters* 88 (1): 109-114.

Fuchs, Thomas, Ludger Wößmann (2007). What Accounts for International Differences in Student Performance? A Re-examination using PISA Data. *Empirical Economics*: forthcoming.

Gundlach, Erich, Ludger Wößmann, Jens Gmelin (2001). The Decline of Schooling Productivity in OECD Countries. *Economic Journal* 111 (471): C135-C147.

Hanushek, Eric A. (2002). Publicly Provided Education. In: Alan J. Auerbach, Martin Feldstein (eds.), *Handbook of Public Economics*, Vol. 4: 2045-2141. Amsterdam: North Holland.

Hanushek, Eric A., Steven G. Rivkin, Lori L. Taylor (1996). Aggregation and the Estimated Effects of School Resources. *Review of Economics and Statistics* 78 (4): 611-627.

Hanushek, Eric A., Ludger Wößmann (2006). Does Early Tracking Affect Educational Inequality and Performance? Differences-in-Differences Evidence across Countries. *Economic Journal* 116 (510): C63-C76.

Jürges, Hendrik, Kerstin Schneider, Felix Büchel (2005a). The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany. *Journal of the European Economic Association* 3 (5): 1134-1155.

Jürges, Hendrik, Wolfram F. Richter, Kerstin Schneider (2005b). Teacher Quality and Incentives: Theoretical and Empirical Effects of Standards on Teacher Quality. *Finanzarchiv* 61 (3): 298-326.

Köller, Olaf, Jürgen Baumert (2001). Leistungsgruppierung in der Sekundarstufe I: Ihre Konsequenzen für die Mathematikleistung und das mathematische Selbstkonzept der Begabung. *Zeitschrift für Pädagogische Psychologie* 15 (2): 99-110.

Kultusministerkonferenz (KMK) (2005). Zentrale oder dezentrale Abiturprüfungen in den Ländern. http://www.kmk.org/schul/Abfrage-Zentralabitur-05-07.pdf [13.2.2007].

Kultusministerkonferenz (KMK) (2006). Übergang von der Grundschule in Schulen des Sekundarbereichs I. http://www.kmk.org/doc/publ/ueberg.pdf [13.2.2007].

Lee, Jong-Wha, Robert J. Barro (2001). Schooling Quality in a Cross-Section of Countries. *Economica* 68 (272): 465-488.

Organisation for Economic Co-operation and Development (OECD) (2004). *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD.

Prenzel, Manfred, Jürgen Baumert, Werner Blum, Rainer Lehmann, Detlev Leutner, Michael Neubrand, Reinhard Pekrun, Jürgen Rost, Ulrich Schiefele (eds.) (2005). *PISA 2003: Der zweite Vergleich der Länder in Deutschland – Was wissen und können Jugendliche?* Münster: Waxmann.

Schütz, Gabriela, Heinrich W. Ursprung, Ludger Wößmann (2005). Education Policy and Equality of Opportunity. CESifo Working Paper 1518. Munich: CESifo.

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

Statistische Ämter der Länder (2006). Volkswirtschaftliche Gesamtrechnungen der Länder: Bruttoinlandsprodukt, Bruttowertschöpfung in den Ländern und Ost-West-Großraumregionen Deutschlands 1991 bis 2005. Reihe 1: Ländererergebnisse, Band 1. Ergebnisse der Revision 2005. Stuttgart: Statistisches Landesamt Baden-Württemberg.

Statistische Ämter des Bundes und der Länder (2006). Internationale Bildungsindikatoren im Ländervergleich. Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt (2004). Kindertagesbetreuung in Deutschland: Einrichtungen, Plätze, Personal und Kosten 1990 bis 2002. Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt (2006a). Bildung und Kultur: Allgemeinbildende Schulen. Fachserie 11, Reihe 1. Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt (2006b). Im Fokus: Ausgaben je Schüler/-in, 2003. Wiesbaden: Statistisches Bundesamt.

Statistisches Bundesamt (2007). Kindertagesbetreuung (3- bis 5-Jährige). Press release, March 8, 2007. Wiesbaden: Statistisches Bundesamt. http://www.destatis.de/presse/deutsch/pm2007/p0990082.htm [8.3.2007].

Toma, Eugenia F. (1996). Public Funding and Private Schooling across Countries. *Journal of Law and Economics* 39 (1): 121-148.

Toma, Eugenia F. (2005). Private Schools in a Global World: 2004 Presidential Address. *Southern Economic Journal* 71 (4): 693-704.

Weiß, Manfred, Corinna Preuschoff (2006). Gibt es einen Privatschuleffekt? Ergebnisse eines Schulleistungsvergleichs auf der Basis von Daten aus PISA-E. In: Manfred Weiß (ed.), *Evidenzbasierte Bildungspolitik: Beiträge der Bildungsökonomie*: 55-71. Schriften des Vereins für Socialpolitik, N.F. Band 313. Berlin: Duncker& Humblot.

Wößmann, Ludger (2002). *Schooling and the Quality of Human Capital*. Berlin: Springer.

Wößmann, Ludger (2003a). Schooling Resources, Educational Institutions and Student Performance: the International Evidence. *Oxford Bulletin of Economics and Statistics* 65 (2): 117-170.

Wößmann, Ludger (2003b). Central Exit Exams and Student Achievement: International Evidence. In: Paul E. Peterson, Martin R. West (eds.), *No Child Left Behind? The Politics and Practice of School Accountability*: 292-323. Washington, D.C.: Brookings Institution Press.

Wößmann, Ludger (2005a). Educational Production in Europe. *Economic Policy* 20 (43): 445-504.

Wößmann, Ludger (2005b). Public-Private Partnerships in Schooling: Cross-Country Evidence on their Effectiveness in Providing Cognitive Skills. PEPG Research Paper 05-09. Cambridge, MA: Harvard University, Program on Education Policy and Governance.

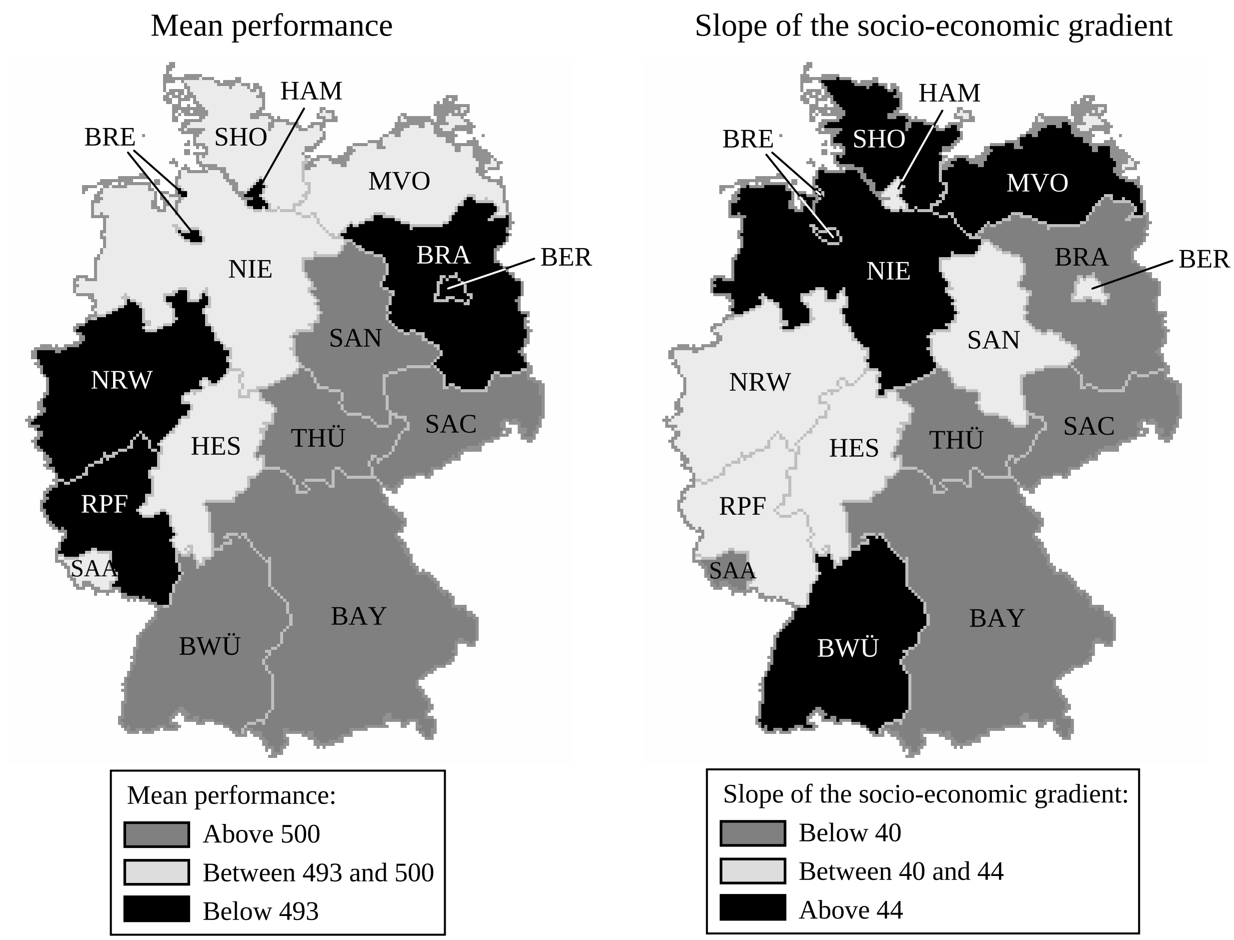
Wößmann, Ludger (2007a). International Evidence on School Competition, Autonomy and Accountability: A Review. *Peabody Journal of Education*: forthcoming.

Wößmann, Ludger (2007b). International Evidence on Expenditures and Class Size: A Review. *Brookings Papers on Education Policy*: forthcoming.

Wößmann, Ludger, Thomas Fuchs, Elke Lüdemann, Gabriela Schütz (in process). School Choice, Autonomy and Accountability: How They Interact to Determine the Efficiency and Equity of Student Achievement. Ifo Institute at the University of Munich.

Wößmann, Ludger, Martin R. West (2006). Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS. *European Economic Review* 50 (3): 695-736.

Figure A1: The German states



Data refer to the PISA-E 2003 math test. See Table 1 for state abbreviations.

Table A1: Schooling data across German states: Definitions, sources, and descriptive statistics

Variable	Definition	Source	Mean	Std. dev.	Min	Max
Mean performance	Mean performance of students on the PISA-E 2003 math test	Prenzel et al. (2005)	498.0	15.4	470.7	532.6
Slope of socio-economic gradient	Slope of the socio-economic gradient (estimated coefficient of a regression of individual student performance on ESCS)	Prenzel et al. (2005)	42.3	3.1	37.5	48.3
External exit exam	External exit exam (dummy)	KMK (2005)	0.44	0.51	0	1
Delayed tracking	Late tracking (dummy equaling 1 if tracking after grade 6 and 0 if tracking after grade 4)	KMK (2006)	0.13	0.34	0	1
Number of tracks	Number of tracks (no. of different types of schools with at least 5 percent of the student population each), not counting the comprehensive schools (captured by next variable)	Statistisches Bundesamt (2006a)	2.88	0.62	2	4
Comprehensive schools	State has integrated comprehensive school type (dummy)	Prenzel et al. (2005)	0.50	0.52	0	1
GDP per capita	Gross domestic product (GDP) per capita, 2003, in 1,000 Euro	Statistische Ämter der Länder (2006)	25.433	7.414	17.777	44.467
Socio-economic background	PISA index of economic, social, and cultural status (ESCS)	Prenzel et al. (2005)	0.16	0.07	0.01	0.26
Standard deviation of socio-economic background	Standard deviation of ESCS	Prenzel et al. (2005)	0.98	0.09	0.85	1.14
Father employed	Share of students whose father has full-time employment	Prenzel et al. (2005)	80.76	3.79	73.0	87.9
Parental education	Longest education of either father or mother, in years	Prenzel et al. (2005)	13.42	0.32	13.0	13.9
Migration background	Share of students with migration background (at least one parent born in a foreign country)	Prenzel et al. (2005)	19.87	11.58	3.6	35.8
City state	State is a city state (dummy)		0.19	0.40	0	1
East German state	State is in east Germany (dummy)		0.38	0.50	0	1
Expenditure per student	Expenditure on public schools per student, 2003, in 1,000 Euro	Stat. Bundesamt (2006b)	5.19	0.54	4.600	6.500
Class size	Average class size in lower secondary education, 2004	Stat. Ämter des Bundes und der Länder (2006)	23.91	1.79	20.3	26.8
Instruction time	Instruction time in math, in hours per week	Prenzel et al. (2005)	3.04	0.17	2.7	3.3
Delayed enrollment	Share of students whose enrollment was delayed	Prenzel et al. (2005)	13.11	3.50	8.0	19.5
Repeaters	Share of students who repeated at least one grade during lower secondary education	Prenzel et al. (2005)	14.66	4.29	7.4	24.7
Pre-school enrollment	Share of children aged 3 to 5 who attend day-care centers or other pre-school facilities, 2005/06	Statistisches Bundesamt (2007)	90.03	5.34	79.7	98.5
Private school enrollment	Share of students attending private schools in lower secondary education, 2004	Stat. Ämter des Bundes und der Länder (2006)	5.99	2.94	2.2	12.2
Conservative voters	Share of votes for CDU/CSU party in total votes for CDU/CSU and SPD in federal Bundestag election of 22.9.2002	Bundeswahlleiter (2002)	0.45	0.08	0.321	0.669
Conservative prime minister	Prime minister of state is from CDU/CSU party in May 2003		0.56	0.51	0	1
Religious affiliation	Share of Protestant (and Catholic) church members in the population, 2003	Evangelische Kirche in Deutschland (2004)	30.60	12.48	15.9	56.9
Transfer by marks	Marks (grades) rather than parents decide on which track a student can go to (dummy)	Demmer (2004)	0.44	0.51	0	1

Table A2: Schooling data across OECD countries: Definitions, sources, and descriptive statistics

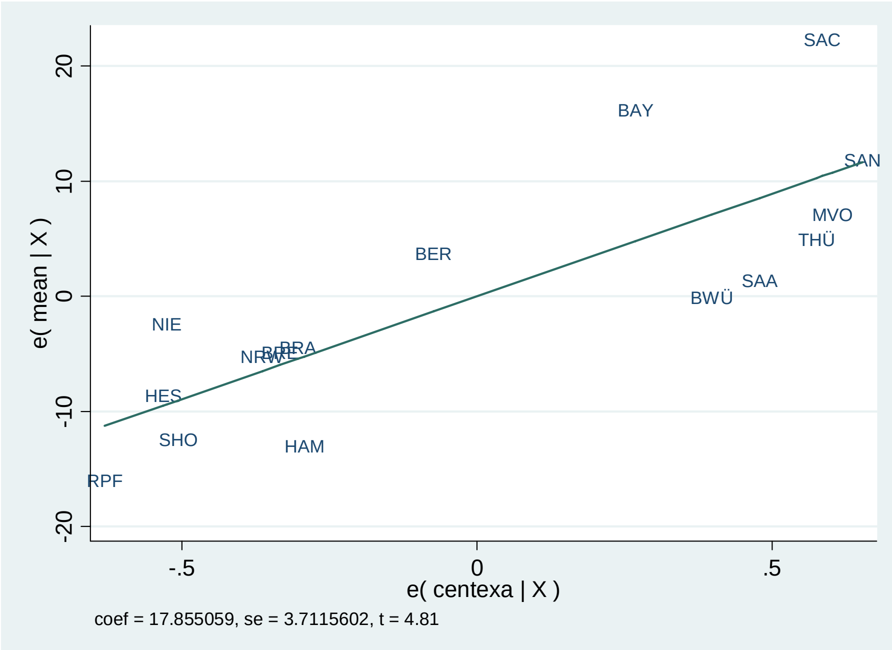
Variable	Definition	Source	Obs.	Mean	Std. dev.	Min	Max
Mean performance	Mean performance of students on the PISA 2003 math test	OECD (2004)	28	499.5	36.6	385	544
Slope of socio-economic gradient	Slope of the socio-economic gradient (estimated coefficient of a regression of individual student performance on ESCS)	OECD (2004)	28	41.5	7.5	28	55
External exit exam	External exit exam (dummy)	Bishop (2006), Fuchs and Wößmann (2007), among others	28	0.658	0.464	0	1
Delayed tracking	Late tracking (dummy equaling 1 if tracking not earlier than age 12)	OECD (2004)	28	0.821	0.390	0	1
Age of first tracking	First age of selection in the education system	OECD (2004)	28	14.0	2.0	10	16
Number of tracks	Number of school types or distinct educational programs available to 15-year-olds	OECD (2004)	27	2.556	1.396	1	5
GDP per capita	Gross domestic product (GDP) per capita, 2003, in 1,000 equivalent US dollars using purchasing power parity	OECD (2004)	27	22.972	8.211	6.046	36.587
Socio-economic background	PISA index of economic, social, and cultural status (ESCS)	OECD (2004)	28	-0.008	0.401	-1.13	0.69
Standard deviation of socio-economic background	Standard deviation of ESCS	OECD (2004)	28	0.914	0.127	0.73	1.27
Father employed	Share of students whose father has full-time employment	PISA student compendium[a]	28	78.5	8.5	51.19	90.82
Expenditure per student	Cumulative expenditure per student between 6 and 15 years, 2003, in equivalent US dollars using purchasing power parity	OECD (2004)	25	52.690	20.271	14.874	79.716
Class size	Average number of students attending math class, reported by PISA 2003 students	OECD (2004)	28	23.4	4.4	17.7	36.1
Private school enrollment	Average share of students attending schools managed directly or indirectly by a non-government organization in PISA 2003	PISA school compendium[a]	26	18.0	22.3	0.45	74.46

[a] Student and school compendia provided by PISA 2003 at http://pisaweb.acer.edu.au/oecd_2003/oecd_pisa_data_s1.html.

Figure 1: Descriptive pattern: Mean performance and slope of the socio-economic gradient

Data refer to the PISA-E 2003 math test. See Table 1 for state abbreviations.

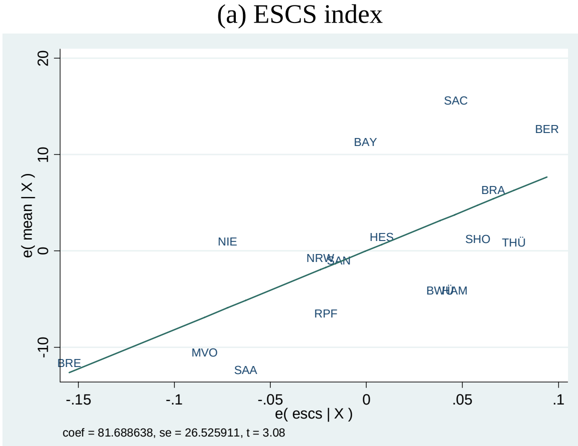
Figure 2: External exit exams and student performance across German states



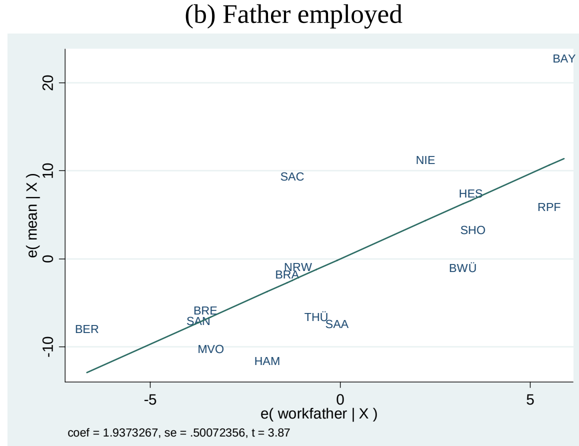
Added-variable plot of student performance on PISA-E 2003 math test against external exit exams, conditional on index of socio-economic background and employment status of father. See specification (1) in Table 2.

Figure 3: Socio-economic background and student performance across German states

(a) ESCS index



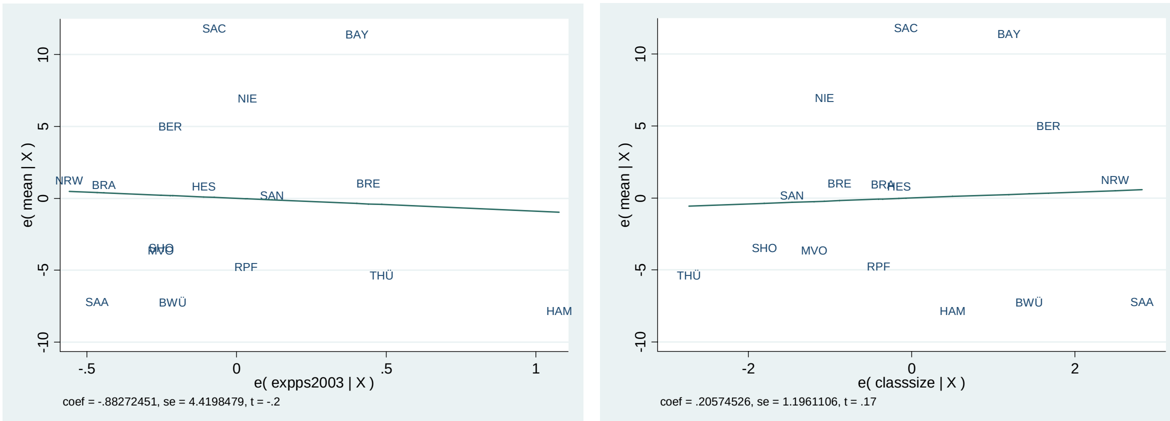
(b) Father employed



Added-variable plots of student performance on PISA-E 2003 math test against index of socio-economic background (ESCS) and share of students with employed father, respectively, conditional on external exit exams and employment status of father/index of socio-economic background. See specification (1) in Table 2.

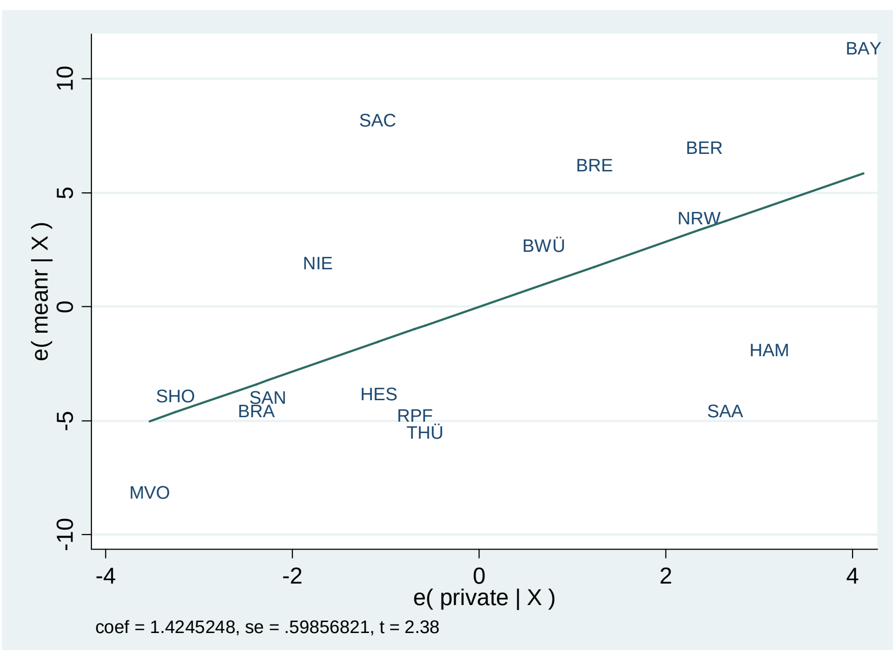
Figure 4: Resources and student performance across German states

(a) Expenditure per student (b) Class size



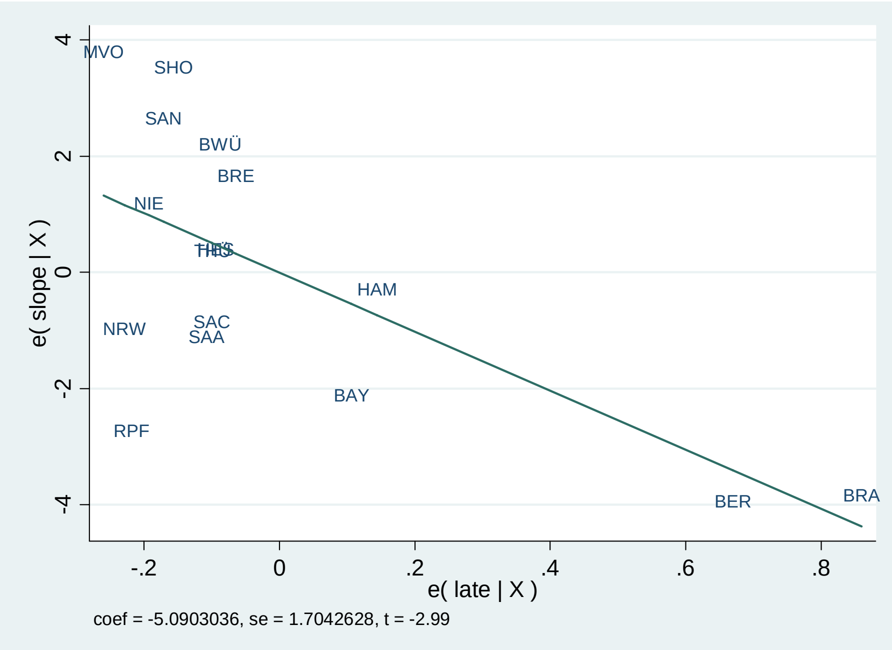
Added-variable plots of student performance on PISA-E 2003 math test against expenditure per student and class size, respectively, conditional on external exit exams, index of socio-economic background, and employment status of father. For plot (a), see specification (5) in Table 2.

Figure 5: Private school enrollment and reading performance across German states



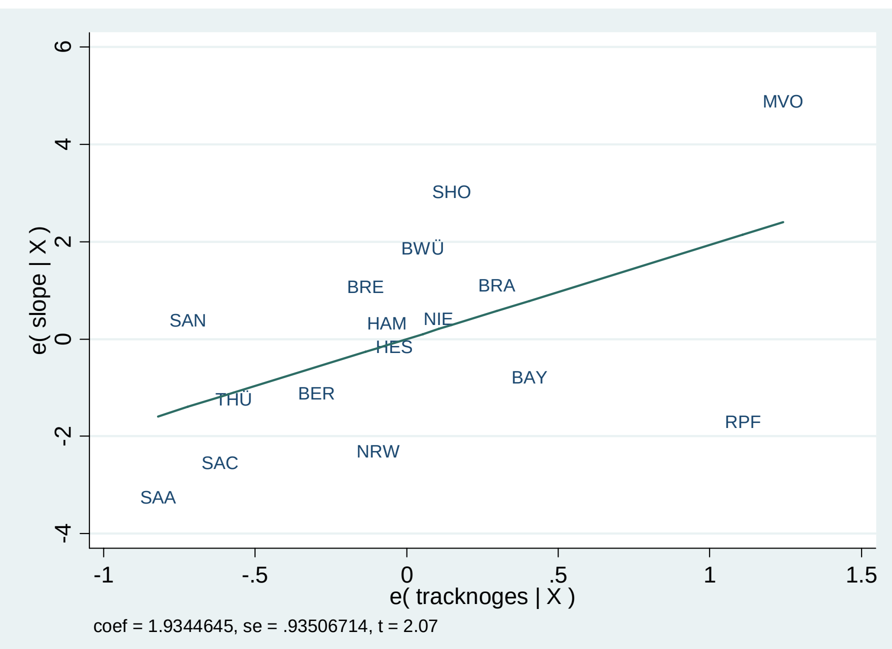
Added-variable plot of student performance on PISA-E 2003 reading test against private school enrollment, conditional on external exit exams, index of socio-economic background. and employment status of father.

Figure 6: Delayed tracking and inequality of opportunity across German states



Added-variable plot of slope of the socio-economic gradient in the PISA-E 2003 math test against delayed tracking, conditional on standard deviation of index of socio-economic background, GDP per capita, and number of tracks. See specification (12) in Table 3.

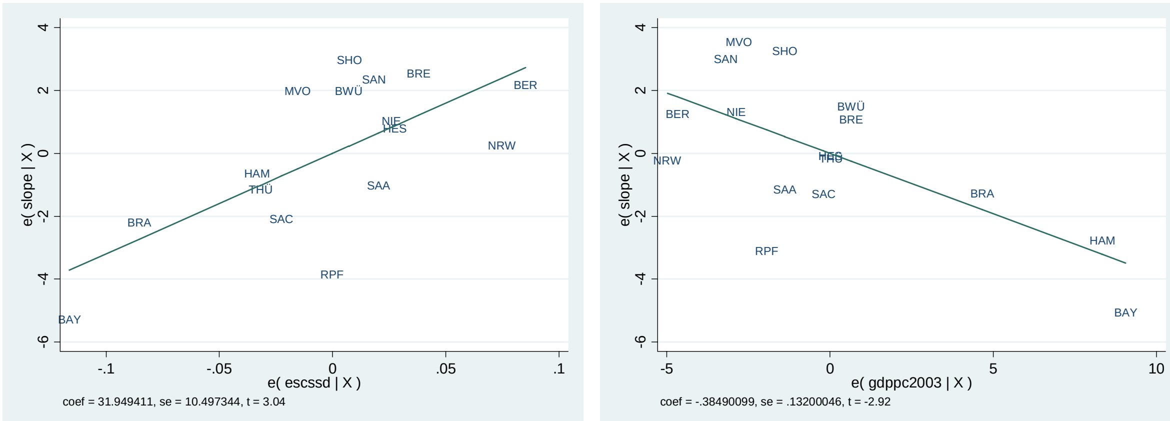
Figure 7: Number of tracks and inequality of opportunity across German states



Added-variable plot of slope of the socio-economic gradient in the PISA-E 2003 math test against number of tracks, conditional on standard deviation of index of socio-economic background, GDP per capita, and delayed tracking. See specification (12) in Table 3.

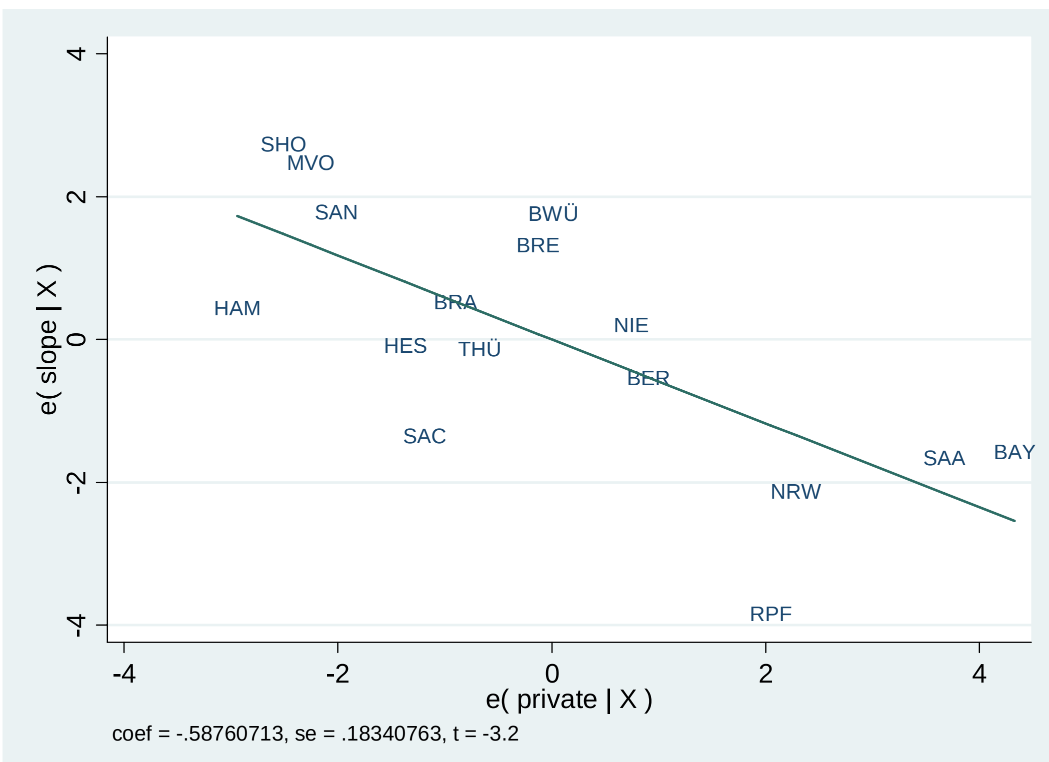
Figure 8: Socio-economic background and inequality of opportunity across German states

(a) Variance of socio-economic background (b) GDP per capita



Added-variable plot of slope of the socio-economic gradient in the PISA-E 2003 math test against standard deviation of index of socio-economic background and GDP per capita, respectively, conditional on delayed tracking, number of tracks, and GDP per capita/standard deviation of socio-economic background. See specification (12) in Table 3.

Figure 9: Private school enrollment and inequality of opportunity across German states



Added-variable plot of slope of the socio-economic gradient in the PISA-E 2003 math test against private school enrollment, conditional on standard deviation of index of socio-economic background, GDP per capita, delayed tracking, and number of tracks. See specification (20) in Table 3.

Table 1: Descriptive state statistics of selected variables

		Mean performance	Slope of socio-economic gradient	External exit exam	Delayed tracking	Number of tracks	Comprehensive schools	GDP per capita	Socio-economic background	Std. dev. of ESCS	Father employed	Migration background	City state	East German state	Expenditure per student	Class size	Instruction time	Private school enrollment
BWÜ	Baden-Wurttemberg	512.0	44.9	1	0	3	0	29.685	0.19	1.03	85.2	31.6	0	0	4.9	25.1	3.3	7.2
BAY	Bavaria	532.6	37.6	1	0	3	0	31.004	0.15	0.92	87.9	20.5	0	0	5.2	25.3	3.1	12.2
BER	Berlin	488.5	41.6	0	1	3	1	23.194	0.26	1.08	73.0	26.1	1	1	5.7	25.0	2.8	4.6
BRA	Brandenburg	492.2	37.5	0	1	3	1	18.348	0.23	0.86	78.3	6.0	0	1	5.0	23.8	2.8	2.2
BRE	Bremen	470.7	45.1	0	0	3	1	35.481	0.01	1.12	76.4	35.8	1	0	5.4	23.2	3.2	8.3
HAM	Hamburg	480.9	41.4	0	0	3	1	44.467	0.21	1.14	77.8	34.6	1	0	6.5	24.6	2.7	7.8
HES	Hessen	496.7	43.6	0	0	3	1	31.621	0.17	1.07	83.2	30.4	0	0	4.8	24.8	3.2	6.2
MVO	Mecklenburg-W. Pomerania	493.2	48.3	1	0	4	0	17.777	0.07	0.91	78.7	4.7	0	1	5.0	21.6	2.8	2.3
NIE	Lower Saxony	494.2	44.4	0	0	3	0	22.932	0.09	0.98	82.1	24.1	0	0	4.8	23.8	3.0	6.3
NRW	North Rhine-Westphalia	486.0	43.4	0	0	3	1	25.985	0.14	1.06	78.7	29.6	0	0	4.6	26.8	3.0	8.4
RPF	Rhineland-Palatinate	492.8	42.2	0	0	4	0	23.147	0.14	0.98	85.3	23.4	0	0	4.7	24.9	3.1	7.8
SAA	Saarland	497.7	39.6	1	0	2	1	23.877	0.09	0.96	82.0	19.9	0	0	4.6	26.1	3.1	9.4
SAC	Saxony	523.2	38.8	1	0	2	0	19.260	0.20	0.87	80.7	5.9	0	1	5.4	22.9	3.2	3.6
SAN	Saxony-Anhalt	502.0	43.3	1	0	2	0	18.190	0.14	0.90	78.3	4.4	0	1	5.6	21.2	3.1	2.4
SHO	Schleswig-Holstein	496.5	46.2	0	0	3	1	23.773	0.22	0.97	83.2	17.3	0	0	4.8	23.1	3.1	3.3
THÜ	Thuringia	509.6	39.8	1	0	2	0	18.194	0.23	0.85	81.3	3.6	0	1	6.0	20.3	3.1	3.9

Means within each state. See Table A1 for definitions and data sources.

Table 2: Mean student performance across German states

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
External exit exam	17.86***	13.25**	15.25**	16.23***	18.11***	17.93***		15.42***	12.76***	19.27***	16.87***
	(3.71)	*(5.36)*	(5.48)	*(4.79)*	*(4.08)*	*(4.53)*		*(4.57)*	*(3.82)*	*(5.98)*	*(3.67)*
Socio-economic background	81.69***	57.67	67.34	77.14**	84.20**	84.65**	82.82**	83.67***	56.71**	88.34**	64.64**
	(26.53)	*(33.48)*	*(38.01)*	*(28.49)*	*(30.37)*	*(29.67)*	*(36.55)*	*(26.78)*	*(23.70)*	*(34.21)*	*(28.65)*
Father employed	1.94***	2.99**	2.07***	2.01***	1.87**	1.77**	1.79**	1.76***	0.49	1.88**	1.88***
	(0.50)	*(1.15)*	*(0.57)*	*(0.53)*	*(0.62)*	*(0.67)*	*(0.74)*	*(0.54)*	*(0.67)*	*(0.63)*	*(0.49)*
City state		3.83									
		(7.75)									
East German state		9.36									
		(7.74)									
GDP per capita			0.002								
			(0.46)								
Parental education			6.72								
			(14.19)								
Migration background				-0.12							
				(0.21)							
Expenditure per student					-0.88						
					(4.42)						
Class size						0.24					
						(1.25)					
Instruction time						5.62					
						(14.34)					
Comprehensive schools							-14.01**	-4.50			
							(5.44)	*(4.88)*			
Conservative voters									89.56**		
									(32.53)		
Conservative prime minister									2.29		
									(3.63)		
Private school enrollment										0.27	
										(0.86)	
Pre-school enrollment										-0.16	
										(0.58)	
Slope of socio-economic gradient											-0.85
											(0.63)
Constant	320.81***	237.15**	223.29	318.96***	330.16***	311.16***	347.04***	338.35***	402.24***	336.63***	364.73***
	(40.34)	*(92.55)*	*(212.12)*	*(41.67)*	*(62.93)*	*(49.86)*	*(61.09)*	*(44.83)*	*(44.65)*	*(63.63)*	*(50.95)*
N	16	16	16	16	16	16	16	16	16	16	16
R^2	0.835	0.856	0.843	0.839	0.835	0.837	0.688	0.846	0.907	0.837	0.858
R^2 (adj.)	0.793	0.784	0.765	0.781	0.775	0.756	0.610	0.791	0.860	0.755	0.806

Dependent variable: student performance on PISA-E 2003 math test. Sample: all German states. Standard errors in parentheses. Significance level: *** 1%, ** 5%, * 10%.

Table 3: Inequality of opportunity across German states

	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
Delayed tracking	-5.09 **	-5.93 ***	-5.33 **	-5.64 **	-6.18 **	-4.41 ***	-4.74 **	-4.82 **	-6.04 ***	-4.74 **	-5.30 **
	(1.70)	*(1.35)*	(2.09)	*(1.88)*	*(2.17)*	*(1.38)*	*(1.69)*	*(1.67)*	*(1.29)*	*(1.93)*	*(1.75)*
Number of tracks	1.93 *	3.21 ***	2.00 *	2.42	2.18 *	2.08 **	1.98 *	1.42	2.02 **	2.08 *	1.81 *
	(0.94)	*(0.84)*	*(1.03)*	*(1.32)*	*(1.01)*	*(0.75)*	*(0.91)*	*(1.00)*	*(0.69)*	*(1.03)*	*(0.96)*
Standard deviation of socio-economic background	31.95 **	31.92 ***	30.77 **	27.04 **	35.45 **	37.85 ***	23.64 *	26.02 **	30.11 ***	34.53 **	24.99 *
	(10.50)	*(8.10)*	*(12.25)*	*(11.90)*	*(14.48)*	*(8.64)*	*(12.28)*	*(11.26)*	*(7.76)*	*(12.31)*	*(13.69)*
GDP per capita	-0.38 **	-0.43 ***	-0.39 **	-0.35 **	-0.33 *	-0.35 ***	-0.32 **	-0.40 **	-0.23 *	-0.39 **	-0.34 **
	(0.13)	*(0.10)*	*(0.14)*	*(0.15)*	*(0.17)*	*(0.11)*	*(0.14)*	*(0.13)*	*(0.11)*	*(0.14)*	*(0.14)*
Comprehensive schools			0.35								
			(1.62)								
Conservative voters				-10.91							
				(8.29)							
Conservative prime minister				0.54							
				(1.88)							
City state					-0.34						
					(3.20)						
East German state					2.14						
					(2.45)						
Class size						-0.80 **					
						(0.29)					
Pre-school enrollment							-0.15				
							(0.12)				
Delayed enrollment								0.31			
								(0.24)			
Private school enrollment									-0.59 ***		
									(0.18)		
External exit exam										0.74	
										(1.64)	
Mean performance											-0.04
											(0.05)
Constant	15.86 *	13.82 **	16.71 *	23.21 *	9.64	27.74 ***	35.66 *	19.60 **	17.12 **	12.68	43.27
	(7.43)	*(5.77)*	*(8.71)*	*(10.60)*	*(14.40)*	*(7.37)*	*(17.65)*	*(7.83)*	*(5.49)*	*(10.43)*	*(34.67)*
N	16	15	16	16	16	16	16	16	16	16	16
R^2	0.669	0.821	0.671	0.724	0.707	0.809	0.713	0.714	0.837	0.676	0.690
R^2 (adj.)	0.549	0.750	0.506	0.540	0.511	0.714	0.569	0.572	0.755	0.514	0.534

Dependent variable: slope of the socio-economic gradient in the PISA-E 2003 math test. Sample: all German states; specification (13) omits one outlier (RPF). Standard errors in parentheses. Significance level: *** 1%, ** 5%, * 10%.

Table 4: Mean student performance across OECD countries and German states

	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
External exit exam	22.48 **	22.48 **	21.70 ***	21.56 ***	18.48 ***	18.55 ***	19.96 ***	20.36 ***
	(10.59)	*(8.85)*	(5.95)	*(6.06)*	*(5.78)*	*(5.48)*	*(7.35)*	*(6.70)*
Socio-economic background	27.71	27.71 *	28.21 **	29.22 **	37.77 ***	38.31 ***	8.87	11.68
	(18.45)	*(15.40)*	*(12.47)*	*(13.51)*	*(12.47)*	*(11.74)*	*(13.05)*	*(12.62)*
Father employed	1.95 **	1.95 ***	1.85 ***	1.85 ***	1.49 **	1.45 ***	2.89 ***	2.64 ***
	(0.83)	*(0.69)*	*(0.57)*	*(0.57)*	*(0.55)*	*(0.52)*	*(0.55)*	*(0.50)*
Private school enrollment					0.51 ***	0.51 ***		0.50 ***
					(0.16)	*(0.15)*		*(0.16)*
Class size				0.19				
				(0.90)				
Germany (dummy)		-11.30			2.51			
		(124.64)			*(11.37)*			
External exit exam x Germany		-4.63						
		(13.52)						
Socio-economic background		53.98						
x Germany		*(74.68)*						
Father employed x Germany		-0.01						
		(1.54)						
Private school enrollment					-0.47			
x Germany					*(1.56)*			
Constant	332.11 ***	332.11 ***	338.01 ***	333.92 ***	360.88 ***	363.57 ***	259.46 ***	272.08 ***
	(67.60)	*(56.45)*	*(44.95)*	*(49.51)*	*(43.66)*	*(41.34)*	*(43.95)*	*(39.34)*
N	28	44	44	44	42	42	54	50
R^2	0.646	0.663	0.648	0.648	0.732	0.732	0.664	0.699
R^2 (adj.)	0.602	0.598	0.622	0.612	0.686	0.703	0.644	0.672
F (Germany and interactions)		0.40			0.05			
Prob. > *F*		0.805			0.953			

Dependent variables: student performance in the PISA-E 2003 math test. Sample: (23) OECD countries (without Germany); (24)-(28): pooled sample of OECD countries and German states; (29)-(30): pooled sample of OECD countries, non-OECD PISA participants, and German states. Standard errors in parentheses. Significance level: *** 1%, ** 5%, * 10%.

Table 5: Inequality of opportunity across OECD countries and German states

	(31)	(32)	(33)	(34)	(35)	(36)
Delayed tracking	-5.87 [g]	-5.87 *	-4.48 *	-6.96 ***	-6.88 **	-10.20 ***
	(3.96)	*(3.31)*	*(2.29)*	*(2.34)*	*(2.57)*	*(2.45)*
Number of tracks	1.71 [g]	1.71 *	1.78 **	1.36 *	1.34	
	(1.03)	*(0.86)*	*(0.76)*	*(0.78)*	*(0.82)*	
Standard deviation of socio-economic background	-28.04 **	-28.04 ***	-30.93 ***	-27.20 ***	-30.51 **	-25.41 ***
	(11.05)	*(9.24)*	*(8.52)*	*(7.38)*	*(11.63)*	*(8.53)*
GDP per capita	-0.01	-0.01	-0.10			0.22 *
	(0.19)	*(0.16)*	*(0.13)*			*(0.12)*
Private school enrollment				0.06	0.07	
				(0.05)	*(0.05)*	
Germany (dummy)		-51.92 **	-44.84 **	-42.87 ***	-43.53 **	-54.61 **
		(20.51)	*(18.65)*	*(15.04)*	*(16.03)*	*(21.47)*
Delayed tracking x Germany		0.78 [g]				
		(5.23)				
Number of tracks x Germany		0.22 [g]				
		(2.38)				
Standard deviation of socio-economic background x Germany		59.99 **	45.16 **	47.46 ***	48.25 ***	71.40 **
		(26.54)	*(19.73)*	*(16.33)*	*(17.64)*	*(27.25)*
GDP per capita x Germany		-0.39 [g]				-0.79 **
		(0.36)				*(0.37)*
Private school enrollment x Germany				-0.90 *	-0.86	
				(0.49)	*(0.51)*	
External exit exam					-0.12 [g]	
					(2.01)	
Mean performance					-0.01 [g]	
					(0.04)	
Constant	67.78 ***	67.78 ***	71.20 ***	67.32 ***	77.66 **	67.24 ***
	(12.58)	*(10.51)*	*(9.69)*	*(6.94)*	*(28.70)*	*(9.08)*
N	26	42	42	42	42	47
R^2	0.496	0.513	0.494	0.551	0.553	0.514
R^2 (adj.)	0.399	0.376	0.407	0.459	0.427	0.441
F (Germany and interactions)		1.47	3.22	3.23	2.95	3.10
Prob. > F		0.226	0.052	0.034	0.048	0.037
F (variables in [g])	4.27	0.42			0.07	
Prob. > F	0.028	0.742			0.932	

Dependent variables: slope of the socio-economic gradient in the PISA-E 2003 math test. Sample: (31) OECD countries (without Germany); (32)-(35): pooled sample of OECD countries and German states; (36): pooled sample of OECD countries, non-OECD PISA participants, and German states. Standard errors in parentheses. Significance level: *** 1%, ** 5%, * 10%.