

ifo Working Papers

Ursachenkomplexe der PISA-Ergebnisse:
Untersuchungen auf Basis der internationalen Mikrodaten

Ludger Wößmann

Ifo Working Paper No. 16

August 2005

Eine elektronische Version ist auf der ifo-Website www.ifo.de zu finden.

Ursachenkomplexe der PISA-Ergebnisse: Untersuchungen auf Basis der internationalen Mikrodaten*

Abstract

The paper estimates the correlates of student performance across countries using student-level data from the PISA-2000 study. There is a substantial relationship of international student performance in math, science, and reading with institutional features. In particular, there are positive associations with accountability through external exit exams, school autonomy in personnel-management and process decisions, and private operation of schools. By contrast, there is no positive association with private school funding, and a negative association with autonomy in areas with scope for decentralized opportunistic behavior. Student performance shows a positive relationship with school autonomy particularly where external exit exams are in place, highlighting the role of external exams as “currency” of school systems. Positive associations of student performance with resource endowments of schools are mainly related to resource quality. While there is a strong association with family background, computer availability at home shows even a negative relationship with student performance on basic skills, after holding other effects constant. The explanatory power of the model between countries is remarkable.

JEL Code: I2.

Keywords: PISA, Bildungsökonomik, Schülerleistungen, Bildungsproduktionsfunktion.

Ludger Wößmann
ifo Institut für Wirtschaftsforschung
an der Universität München und CESifo
Poschingerstr. 5
81679 München
Tel.: 089/9224 1699
E-Mail: woessmann@ifo.de
Internet: www.cesifo.de/link/woessmann_1.htm

* Der Autor dankt den Teilnehmern der Jahrestagung der Sektion International und Interkulturell Vergleichende Erziehungswissenschaft der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE) am 2.-4. März 2005 in Münster zum Thema „Erziehungswissenschaftliche und politische Relevanz und Rezeption international vergleichender Schulleistungsuntersuchungen“ für zahlreiche hilfreiche Kommentare und Diskussionen.

1. Einleitung: Eine bildungsökonomische Perspektive zu PISA

Die PISA-Studie (Programme for International Student Assessment) ist – in Deutschland noch mehr als anderswo – in aller Munde. Und das zurecht, lässt sich aus bildungsökonomischer Sicht hinzufügen. Denn umfangreiche Studien belegen, dass ein empirisch gesicherter positiver Zusammenhang zwischen den in internationalen Vergleichstests gemessenen Bildungsleistungen und der langfristigen wirtschaftlichen Entwicklung von Volkswirtschaften besteht.¹ Dementsprechend ist es nur natürlich, dass sich die bildungsökonomische Forschung in den letzten Jahren zunehmend dafür interessiert, worin denn die wichtigsten Ursachen für das Abschneiden in den internationalen Schülerleistungstests bestehen. Dabei werden zumeist drei große Ursachenkomplexe betrachtet: der familiäre Hintergrund der Schüler, die Ressourcenausstattung der Schulen und die institutionellen Rahmenbedingungen des Schulsystems.

Die vorliegende Studie berichtet über umfangreiche Arbeiten, die sich dieser Fragestellung anhand der internationalen Daten der PISA-Studie von 2000 empirisch nähern. Dazu werden die Mikrodaten über die Leistungen der einzelnen Schüler in Mathematik, Naturwissenschaften und Lesen mit Daten aus Hintergrundfragebögen der Schüler und Schulen verbunden, welche detaillierte Informationen zu den drei untersuchten Ursachenkomplexen liefern. Eine multivariate Herangehensweise ermöglicht es dabei, bei der Schätzung des Zusammenhangs der Schülerleistungen mit jedem einzelnen Einflussfaktor die Einflüsse aller weiteren Einflussfaktoren konstant zu halten (vgl. Wößmann 2003e, Abschnitt III). Das heißt, dass etwa der Zusammenhang mit der Schulautonomie als einem institutionellen Einflussfaktor so berechnet wird, als wenn alle Schulen den gleichen familiären Hintergrund aufweisen würden. Damit kann eine Konfundierung der jeweiligen Effekte mit anderen beobachteten Hintergrundeffekten ausgeschlossen werden. Die Berücksichtigung der internationalen Variation in den Schülerleistungen ermöglicht es dann, die Verbindung zwischen Schülerleistungen und institutionellen Unterschieden der Schulsysteme, die innerhalb einzelner Länder oftmals gar nicht gegeben sind, zu untersuchen. Im vorliegenden Beitrag kann es aufgrund der Komplexität der Untersuchungen nur um eine kurze Einführung und einen ersten Überblick über die zentralen Befunde der Arbeiten gehen. Details zu methodischen Fragen, weiteren Ergebnis-

¹ Vgl. Hanushek und Kimko (2000), Barro (2001), Wößmann (2002, 2003d) und Gundlach und Wößmann (2004).

sen, theoretischen Grundlagen und Interpretationen sowie ausführlichere Literaturverweise finden sich in Fuchs und Wößmann (2004a, 2004b) und Wößmann (2005a).²

Aus bildungspolitischer Sicht besteht ein zentraler Befund der vorgestellten Arbeiten darin, dass die internationalen Unterschiede in den Schülerleistungen in systematischer Weise mit zahlreichen institutionellen Rahmenbedingungen des Schulsystems variieren. Die ökonomische Theorie legt nahe, dass Institutionen wie zentrale Prüfungssysteme, Autonomie der Schulen und Wettbewerb durch private Anbieter den Erfolg des Bildungsprozesses signifikant beeinflussen können, weil sie die Anreize der beteiligten Personen bestimmen.³ Entsprechend dieser theoretischen Basis belegen die hier vorgestellten empirischen Ergebnisse etwa, dass individuelle Schülerleistungen besser sind in Ländern, in denen zentrale Abschlussprüfungen durchgeführt werden, in Schulen, die Autonomie im Personalmanagement und in Prozessentscheidungen haben, und in Schulen, die in privater Trägerschaft organisiert sind. Demgegenüber zeigt sich kein eindeutiger Zusammenhang der Schülerleistungen mit dem Anteil privater Finanzierung der Schule, und ein negativer Zusammenhang mit Schulautonomie in Entscheidungsbereichen mit Raum für dezentrales opportunistisches Verhalten. Schließlich erweisen sich zentrale Abschlussprüfungen als „Währung“ des Schulsystems, die sicherstellt, dass ein dezentral organisiertes System autonomer Schulen zu hohen Schülerleistungen führt. Diese Befunde eines starken systematischen Zusammenhangs zwischen Schülerleistungen und den Institutionen des Schulsystems legen nahe, dass die Schulpolitik durch die institutionelle Gestaltung des Schulsystems durchaus weite Spielräume hat, auf die Leistungen der Schüler Einfluss zu nehmen.

Die Ergebnisse bestätigen ganz ähnliche Befunde von internationalen Mikroanalysen auf Schülerebene anhand früherer internationaler Schülerleistungstests (vgl. etwa Wößmann 2001b, 2003a, 2003b). Als empirische Analysen möglicher Ursachen von Schülerleistungen sind diese Untersuchungen in der bildungsökonomischen Literatur so genannter „Bildungsproduktionsfunktionen“ beheimatet, zu der vor allem unzählige nationale Studien auf Schüler- oder Schulebene zählen (vgl. Hanushek 2002 für einen Überblick). Die wenigen vorliegenden bildungsökonomischen Analysen von internationalen Leistungstests wurden zunächst nur auf der Ebene der Teilnehmerstaaten durchgeführt (vgl. Bishop 1997, 2005; Hanushek und

² Weitere methodische Details sowie Ergebnisse auf Basis weiterer internationaler Schülerleistungstests (TIMSS, TIMSS-Repeat und IGLU) finden sich in Wößmann (2001a, 2001b, 2003a, 2003b, 2003c, 2003e) und Wößmann und Fuchs (2005).

³ Vgl. Bishop und Wößmann (2004), Wößmann (2004a) und die Literaturverweise in Fuchs und Wößmann (2004a) für Details zur theoretischen Fundierung der hier analysierten Institutionen.

Kimko 2000; Lee und Barro 2001). Demgegenüber können die hier präsentierten Mikrostudien lokale Unterschiede in den Hintergründen der Schüler wesentlich besser berücksichtigen und darüber hinaus noch mögliche Heterogenitäten systemischer Einflussfaktoren innerhalb der Schulsysteme abbilden.

An diese Einleitung schließt in Abschnitt 2 zunächst eine Diskussion der verwendeten Methodik sowie der Daten an. Abschnitt 3 berichtet dann die zentralen Ergebnisse der Untersuchungen in Bezug auf die drei untersuchten Ursachenkomplexe familiärer Hintergrund, schulische Ressourcenausstattung und institutionelle Rahmenbedingungen. Dabei legt ein kurzer Exkurs am Beispiel des Zusammenhangs zwischen Computern und Schülerleistungen die Notwendigkeit multivariater Vorgehensweise dar. Ein weiteres besonderes Augenmerk liegt auf Interaktionseffekten zwischen externen Abschlussprüfungen und Schulautonomie. Der Abschnitt schließt mit Ergebnissen zum Erklärungsgehalt der geschätzten Modelle für die Variation zwischen den Teilnehmerstaaten. Abschnitt 4 liefert einige Schlussbemerkungen.

2. Methodik und Daten

2.1 Methodik: Multivariate Regressionsanalyse internationaler „Bildungsproduktionsfunktionen“

Als empirisches Grundmodell liegt den hier berichteten Untersuchungen folgende Schätzgleichung zugrunde:

$$T_{is} = a + F_{is}\beta_1 + R_{is}\beta_2 + I_s\beta_3 + P_l\beta_4 + \varepsilon_{is} \quad (1)$$

mit T_{is} als Testleistung des Schülers i in Schule s , F als Vektor von Variablen des familiären Hintergrunds (was hier Schüler- und Familienmerkmale sowie häusliche Inputs umfasst), R als Vektor von Maßen der Ressourcenausstattung der Schulen, I als Vektor von Maßen der institutionellen Gegebenheiten der Bildungssysteme und P_l als zentrale Abschlussprüfungen in Land l (als eine vom Vektor I separat betrachtete Institution). Zusätzlich berücksichtigt das Schätzmodell eine Konstante a sowie einen Fehlerterm ε , dessen Eigenschaften im Folgenden noch näher besprochen werden.⁴

⁴ Zur Behandlung von in den Antwortfragebögen fehlenden Werten wurde eine auf geordneten Probit-Schätzungen basierende Imputationsmethode angewandt. Deshalb beinhalten die Schätzgleichungen als zusätzliche Kontrollvariablen Dummies für imputierte Werte für jede im Modell enthaltene Variable sowie für Interaktionen dieser Dummies mit der jeweiligen Variable, was der Vermeidung von Verzerrungen aufgrund fehlender Werte beitragen soll; siehe Fuchs und Wößmann (2004a) für Details.

Die Parametervektoren β_1 bis β_4 , die die Stärke des Zusammenhangs zwischen den Testleistungen und den jeweiligen Einflussfaktoren widerspiegeln, werden in einer mikroökonomischen Regressionsanalyse empirisch geschätzt.⁵ Dabei werden die Daten aller an der PISA-Studie teilnehmenden Schüler aus allen Teilnehmerstaaten in einer Schätzung berücksichtigt. Damit lässt sich die Schätzgleichung als internationale „Bildungsproduktionsfunktion“ bezeichnen, welche einen „Output“ des Bildungsprozesses (die gemessenen Testleistungen) den verschiedenen „Inputs“ gegenüberstellt.⁶

Die gewählte Spezifikation der Schätzgleichung ist zunächst sehr restriktiv, indem sie den Einfluss jeder Variablen so restringiert, dass er in allen Ländern und auf allen Ebenen (innerhalb von Schulen, zwischen Schulen und zwischen Ländern) einheitlich ist. Als solches können die geschätzten Ergebnisse nicht beanspruchen, das heterogene Geflecht tatsächlicher Einflussfaktoren widerzuspiegeln, da zu vermuten ist, dass diese nicht notwendigerweise in allen Situationen einheitlich sind. Mit der restriktiven Spezifikation „erkauft“ man sich aber andererseits den Vorteil, dass die unüberschaubare Informationsflut der PISA-Daten auf ein informatives Maß reduziert werden kann, welches generelle Muster („patterns“) möglicher Wirkungszusammenhänge widerspiegeln kann. Obwohl es selbstverständlich von großem Interesse sein kann, die mögliche Heterogenität verschiedener Wirkungszusammenhänge zwischen Ländern und Ebenen zu analysieren, scheint es für die Fragestellung der vorliegenden Untersuchung – mögliche Bestimmungsgründe der internationalen PISA-Unterschiede – angebracht, von dieser Heterogenität weitgehend zu abstrahieren und statt dessen einen einheitlichen „Durchschnittseffekt“ für jede Variable zu schätzen. In Wößmann (2003a) wird die hier gewählte restriktive Spezifikation mit einer alternativen zweistufigen Spezifikation verglichen, welche im ersten Schritt eine Analyse innerhalb der Länder vollführt, um dann im zweiten Schritt die verbleibende Variation zwischen den Ländern zu untersuchen. Es zeigt sich, dass sich die substantiellen Ergebnisse der beiden Modellspezifikationen nicht unterscheiden. Wie weiter unten noch erläutert wird, berichten wir hier aber auch noch Ergebnisse einer erweiterten Modellspezifikation, die die Restriktivität der Einheitlichkeitsannahme mildert, indem sie systematische Heterogenität in einigen Einflüssen zulässt.

Bei der Schätzung der Regressionsgleichungen ist zu berücksichtigen, dass es sich bei den PISA-Daten um Erhebungsdaten mit einer komplexen Datenstruktur handelt. Dabei verlangt

⁵ Für eine Einführung in die Grundmethodik der Regressionsanalyse vgl. jedes einschlägige Einführungslehrbuch in die Ökonometrie, z.B. Stock und Watson (2003).

⁶ Siehe etwa Hanushek (2002) für Details zur Spezifikation von Bildungsproduktionsfunktionen.

die hierarchische Struktur der Daten nach einer mehrebenenanalytischen Vorgehensweise. Diesem Tatbestand wird in den vorliegenden Untersuchungen dadurch Rechnung getragen, dass im Rahmen von Cluster-robusten linearen Regressionen (CRLR) eine nicht-triviale Struktur der Fehlerterme abgebildet wird. Da Schulen in PISA die primäre Erhebungseinheit darstellen und da zahlreiche Variablen auf der Ebene der Schulen erhoben werden, so dass die individuellen Observationen innerhalb jeder Schule nicht als unabhängig voneinander betrachtet werden können, hat der Fehlerterm ε nicht nur die übliche Komponente auf der Ebene des individuellen Schülers ν_i , sondern zusätzliche eine Komponente auf Schulebene η_s :

$$\varepsilon_{is} = \eta_s + \nu_i \quad (2)$$

CRLR erlaubt daher bei der Berechnung der statistischen Signifikanz der Schätzergebnisse jegliche Kovarianz der individuellen Observationen innerhalb jedes „Clusters“, d.h. innerhalb jeder Schule. Dies wird implementiert, indem eine Block-diagonale Varianz-Kovarianz-Matrix der Fehlerterme mit einem Block für jede Schule, welcher der tatsächlich observierten Kovarianzstruktur der Kleinstquadrat-Fehlerterme entspricht, zugelassen wird (für Details vgl. Moulton 1986; Deaton 1997, S. 74-78).⁷ Diese Vorgehensweise liefert selbst bei arbiträrer Korrelation innerhalb der primären Erhebungseinheiten und bei zwischen den primären Erhebungseinheiten variierenden Fehlervarianzen eine konsistente Schätzung der Varianz-Kovarianz-Matrix (White 1984, S. 134-142). Damit lockert CRLR die klassische Annahme der Unabhängigkeit aller Observationen dahingehend, dass Unabhängigkeit nur zwischen den primären Erhebungseinheiten gewährleistet sein muss.⁸

Zusätzlich zu dieser Mehrebenenstruktur der Daten ist zu berücksichtigen, dass PISA ein stratifiziertes Erhebungsdesign innerhalb jeden Landes genutzt hat, was zu variierenden Erhebungswahrscheinlichkeiten führt. Um für die Grundgesamtheit repräsentative Schätzungen der Koeffizienten zu erhalten, wird in den Regressionen daher jeder Schüler mit seiner Erhebungswahrscheinlichkeit in dem jeweiligen Land gewichtet, so dass jede Stratifizierungsein-

⁷ Bei Variablen, die nur auf der Ebene der Teilnehmerländer variieren, wird sogar jegliche vorhandene Kovarianz innerhalb der Länder in den Fehlertermen berücksichtigt.

⁸ In praktischer Hinsicht kommt die in der Ökonometrie häufiger verwendete Cluster-robuste lineare Regression (CRLR) zumeist zu qualitativ identischen Ergebnissen wie die verwandte, in der Erziehungswissenschaft häufiger verwendete Methode der Hierarchical Linear Models (HLM; vgl. etwa Bryk und Raudenbush 1992). Cohen und Baldi (1998) haben allerdings gezeigt, dass der Vorteil der CRLR, nicht der strikten HLM-Annahmen von zufällig und normal verteilten Effekten zu bedürfen, dazu führt, dass CRLR auch dann sowohl konsistente als auch robuste Parameter und Standardfehler liefert, wenn diese Annahmen durch Ausreißer oder eine verzerrte Fehlerstruktur verletzt werden, während HLM in diesen Fällen zu signifikant unterschätzten Standardfehlern auf höheren Ebenen bzw. verzerrten Parameterschätzungen führen kann.

heit mit einer zu ihrem Anteil an der Gesamtpopulation proportionalen Beitrag in die Schätzung eingeht (vgl. DuMouchel und Duncan 1983; Wooldridge 2001). In die internationale Schätzung gehen dann alle Teilnehmerstaaten mit der gleichen Gewichtung ein.

Da Schülerleistungen offensichtlich von vielen Faktoren abhängen, werden in den Regressionsanalysen viele mögliche Bestimmungsfaktoren gleichzeitig berücksichtigt. Diese multivariate Vorgehensweise dient dazu, Konfundierungen der tatsächlichen Einflussfaktoren mit anderen Faktoren möglichst weitgehend auszuschließen, indem die Einflüsse dieser anderen Faktoren in der Schätzung kontrolliert, also „konstant gehalten“ werden. Dies ist aus methodischer Sicht in dem Maße möglich, wie die Schätzungen trotz miteinander korrelierter Einflussfaktoren zu statistisch signifikanten Ergebnissen gelangen. Zu welcher fehlgeleiteten Schlussfolgerungen eine bivariate Analyse führen kann, die etwa nicht für Einflüsse des familiären Hintergrunds kontrolliert, wird weiter unten am Beispiel des Einflusses von Computern auf die Schülerleistungen noch dargestellt (vgl. Abschnitt 3.2).

Die Auswahl der analysierten Ursachenkomplexe und der darin betrachteten Variablen basiert in weiten Teilen auf einer ökonomisch geleiteten Hypothesenauswahl. Dabei werden einige aus pädagogischer Sicht höchst interessante Einflussfaktoren wie etwa die Lehrmethoden oder die Motivation der Schüler und des Lehrpersonals nicht explizit berücksichtigt. Der methodische Grund hierfür liegt insbesondere darin, dass Lehrmethoden und Motivation in Bezug auf Bildungsprozess, Lehrerleistung und Schülerleistung in höchstem Maße endogene Eigenschaften sein dürften. In diesem Sinne ließen sie sich eher als „Output“ denn als „Input“ des Bildungsprozesses betrachten. Würde man sie statt dessen als „Input“ berücksichtigen, käme es aufgrund ihrer Endogenität zu Verzerrungen der geschätzten Koeffizienten.

Ein ähnliches Problem der Endogenität besteht auch bei zahlreichen Maßen der Ressourcenausstattung der Schulen, zumindest bei Variationen innerhalb der Teilnehmerländer. Oftmals werden Ressourcen wie Klassengrößen, Unterrichtsmaterialien oder Unterrichtszeit direkt aufgrund der (vorhergehenden) Leistungen der Schüler bestimmt. Solche systematischen, endogenen Zuordnungen von Ressourcen auf Schüler verzerren die geschätzten Koeffizienten in den lediglich in einem Querschnitt vorliegenden PISA-Daten. Ein Weg, zumindest Verzerrungen aufgrund der systematischen Zuordnung von Klassengrößen innerhalb von Schulen zu vermeiden, besteht darin, die tatsächliche Klassengröße mit dem Schüler-Lehrer-Verhältnis der Schule in einer zweistufigen Kleinstquadrateschätzung zu instrumentieren (Akerhielm 1995), was in den berichteten Untersuchungen gemacht wurde. Dies stellt sicher, dass nicht-zufällige Variationen innerhalb der Schulen die geschätzten Koeffizienten nicht beeinflussen.

Allerdings unterliegt diese Schätzmethode immer noch dem Problem möglicher systematischer Zuordnungen zwischen Schulen, was durchaus das gravierendere Endogenitätsproblem darstellen könnte (siehe West und Wößmann 2005).⁹ Außerdem ist eine ähnliche Vorgehensweise bei den anderen Ressourcenvariablen aufgrund der gegebenen Datenlage nicht umsetzbar. Die PISA-Daten bieten keine offensichtliche Möglichkeit, die verbleibenden Endogenitätsverzerrungen zu bewältigen. Daher können die vorliegenden Ergebnisse auch nicht notwendigerweise in kausaler Weise interpretiert werden, sondern sind als deskriptive Befunde zu bewerten (vgl. Fuchs und Wößmann 2004a, 2004b für weitere Details).¹⁰

Wie bereits angedeutet, lässt sich die restriktive Annahme der in allen Schulsystemen einheitlichen Effekte in einer erweiterten Spezifikation abschwächen. So werden in Abschnitt 3.5 die Ergebnisse einer Spezifikation berichtet, die es erlaubt, dass alle Effekte zwischen Systemen mit externen Abschlussprüfungen und solchen ohne externe Abschlussprüfungen variieren. Dazu werden in der Modellspezifikation (1) zusätzlich Interaktionseffekte zwischen externen Abschlussprüfungen und den anderen Einflussfaktoren berücksichtigt:

$$T_{is} = a + F_{is}\beta_1 + R_{is}\beta_2 + I_s\beta_3 + P_l\beta_4 + (P_l F_{is})\beta_5 + (P_l R_{is})\beta_6 + (P_l I_s)\beta_7 + \varepsilon_{is} \quad (3)$$

Eine solche Spezifikation erlaubt systematische Heterogenität in den geschätzten Zusammenhängen, was sich in den berichteten Ergebnissen zumindest in Bezug auf Interaktionseffekte zwischen externen Abschlussprüfungen und Schulautonomie als zur Abbildung der Realität sehr wichtige Modellerweiterung erweist.

2.2 Daten: Die internationale PISA-Datenbank

Die PISA-Studie, durchgeführt von der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) im Jahre 2000, testete Schülerleistungen in den drei Kompetenzbereichen Lesekompetenz, mathematische Grundbildung und naturwissenschaftliche Grundbildung. Mit diesen Basiskompetenzen wird natürlich nur *ein* „Outputbündel“ des Bildungsprozesses abgebildet, und es sind sich wohl alle Beobachter einig, dass es noch weitere wichtige Bildungsziele innerhalb und außerhalb der Lehrpläne gibt. Allerdings wird wohl auch kein Beobachter diesen Basiskompetenzen jegliche Bedeutung absprechen wollen, und wie in der Einleitung berichtet haben sie sich als wirtschaftlich höchst relevant erwiesen. Gleichzeitig

⁹ Zusätzliche Verzerrungen der anhand der PISA-Daten geschätzten Ressourcenkoeffizienten können aufgrund von Aggregationsverzerrungen (Fertig und Wright 2005) und aufgrund von über die Zeit variierenden Ausstattungsniveaus (vgl. Gundlach und Wößmann 2000; Gundlach et al. 2001) auftreten.

¹⁰ Wößmann (2005c) und Wößmann und West (2005) untersuchen kausale Klassengrößeneffekte mit spezifischen Schätzmethoden anhand der TIMSS-Daten.

erscheint es auch als wenig plausibel, dass die verschiedenen Bildungsleistungen negativ korrelieren; erst dann wäre ja ein möglicher Zielkonflikt zu berücksichtigen.

Die PISA-Studie wurde 2000 in 32 Ländern durchgeführt, darunter 28 wirtschaftlich entwickelte OECD-Länder.¹¹ Die Zielgruppe des Tests waren 15-jährige Schüler, unter denen in jedem Teilnehmerstaat eine repräsentative Stichprobe erhoben wurde. Die vorliegenden Untersuchungen nutzen die individuellen Schülerdaten, welche für 174.227 Schüler (in 6.626 Schulen) in Lesen, 96.855 Schüler in Mathematik und 96.758 Schüler in Naturwissenschaften vorliegen. Der Test bestand aus lebensnahen Fragestellungen, die solche Basiskompetenzen überprüfen sollten, welche in modernen Staaten für eine Teilnahme am gesellschaftlichen, wirtschaftlichen und politischen Leben notwendig sind. Mithilfe eines auf der Item Response Theory (IRT) basierenden Skalierungsverfahrens wurden die Testergebnisse auf einer internationalen Skala mit einem Mittelwert von 500 Punkten und einer Standardabweichung von 100 Punkten (jeweils in Bezug auf die OECD-Länder) abgebildet.

Neben den Testleistungsdaten liefert die internationale PISA-Datenbank umfangreiche Hintergrundinformationen über den familiären Hintergrund aus Schüler-Hintergrundfragebögen, über die schulische Ressourcenausstattung aus Schüler- und Schulleiter-Hintergrundfragebögen sowie über institutionelle Gegebenheiten des Schulsystems aus Schulleiter-Hintergrundfragebögen. Detaillierte Informationen über die PISA-Studie finden sich in den offiziellen OECD-Publikation (OECD 2000, 2001, 2002; Adams und Wu 2002) sowie unter <http://www.pisa.oecd.org>. Weitere Informationen und internationale deskriptive Statistiken zu dem in den hier berichteten Untersuchungen verwendeten Datensatz werden in Fuchs und Wößmann (2004a) berichtet.

3. Ergebnisse

Die Ergebnisse der Schätzung des empirischen Grundmodells in den drei Kompetenzbereichen werden in Tabelle 1 berichtet. Die wichtigsten Ergebnisse werden im Folgenden nur beispielhaft innerhalb der drei Ursachenkomplexe familiärer Hintergrund, schulische Ressourcenausstattung und institutionelle Rahmenbedingungen angesprochen.

¹¹ In den berichteten Untersuchungen werden die Daten für Liechtenstein aufgrund mangelnder international vergleichbarer Rahmendaten nicht berücksichtigt.

3.1 Schülereigenschaften, familiärer Hintergrund und häusliche Inputs

Neben der jeweiligen Jahrgangsstufe und dem Alter der Schüler wird in den Schätzungen auch das Geschlecht berücksichtigt.¹² Dabei zeigt sich, dass Mädchen in der Lesekompetenz im Durchschnitt besser abschneiden als Jungen, während es in Mathematik und Naturwissenschaften umgekehrt ist. Da die Testergebnisse so skaliert sind, dass sie eine internationale Standardabweichung von 100 Punkten haben, lassen sich die Größen aller geschätzten Koeffizienten als Prozentpunkte einer internationalen Standardabweichung interpretieren. Somit schneiden Mädchen im Lesen 23,7 Prozent einer internationalen Standardabweichung besser ab als Jungen. Als konkreten Vergleichspunkt für die Effektgrößen lässt sich auch die (unkonditionierte) Leistungsdifferenz zwischen der 9. und 10. Klasse (den beiden Jahrgangsstufen mit dem größten Anteil 15-Jähriger) heranziehen, welche in Mathematik 30,3, in Naturwissenschaften 32,4 und in Lesen 33,2 Punkte beträgt. Dementsprechend entspricht der Leistungsvorsprung der Mädchen in Lesen gut zwei Drittel dieses Jahrgangsstufenäquivalents.

Zahlreiche Maße des familiären Hintergrunds der Schüler weisen einen statistisch signifikanten und quantitativ bedeutsamen Zusammenhang mit den PISA-Leistungen auf. Dies trifft auf den Immigrationsstatus der Schüler und der Eltern (vgl. dazu auch Entorf und Minoiu 2005), den Familienstatus,¹³ den Arbeitsstatus und Beruf der Eltern sowie (in geringerem Maße) die geographische Lage zu. Den stärksten Zusammenhang weisen aber die beiden Maße „Bildungsstand der Eltern“ und „Bücher im Haushalt“ auf, die beide als Indikatoren des sozialen, ökonomischen und Bildungs-Hintergrunds der Eltern dienen können. Damit erweist sich der familiäre Hintergrund als von zentraler Bedeutung für die erzielten Fähigkeiten der Schüler, der die meisten schulischen Einflüsse quantitativ weit in den Schatten stellt. Tendenziell erweisen sich die Zusammenhänge der Leseleistungen mit dem familiären Hintergrund als etwas größer als die der Mathematik- und Naturwissenschaftsleistungen.

Signifikant positive Zusammenhänge der Schülerleistungen ergeben sich auch mit den häuslichen Inputs der elterlichen Unterstützung (in der Einschätzung der Schulleiter) und der Dauer der Hausarbeiten, wobei der Zusammenhang im Falle der Leseleistungen nicht linear ansteigend ist.

¹² Da PISA eine altersbasierte Zielgruppe hat, könnte die Jahrgangsstufe in Systemen mit Klassenwiederholung zu einem gewissen Grade endogen in Bezug auf die Schülerleistungen sein, was gegen eine Aufnahme als Kontrollvariable spräche. Allerdings kontrolliert die Jahrgangsstufe in der internationalen Schätzung auch für Unterschiede im Einschulungsalter. Deshalb wurden die vorliegenden Schätzungen auch auf Robustheit gegen ein Auslassen der Jahrgangsstufenvariablen getestet, wobei sich die Ergebnisse als qualitativ nicht sensitiv erwiesen haben.

¹³ Für eine Analyse von Effekten der Geschlechterrivalität in PISA vgl. Wolter und Vellacott (2003).

3.2 Computer und Schülerleistungen: Ein Exkurs über die Notwendigkeit multivariater Vorgehensweise

Als weiterer häuslicher Input weist die Verfügbarkeit von Computern im Haushalt in wohl unerwarteter Weise einen signifikant negativen Zusammenhang mit den Schülerleistungen in allen drei Kompetenzbereichen auf. Interessanterweise zeigt sich dieser negative Zusammenhang aber nicht bei einer rein bivariaten Betrachtungsweise. Betrachtet man nämlich lediglich den einfachen Zusammenhang zwischen Computerverfügbarkeit zu Hause und PISA-Leistungen, wie in Spalte I in Tabelle 2 beispielhaft für den Fall der Mathematikleistungen berichtet,¹⁴ so findet sich ein starker signifikanter positiver Zusammenhang.

Doch dieses bivariate Ergebnis ist trügerisch, wie sich bei der zusätzlichen Berücksichtigung weiterer Einflussfaktoren erweist. So wird der Zusammenhang schon schwächer, wenn man die Einflüsse der Variablen der Schülereigenschaften (vgl. Tabelle 1) herausrechnet (Spalte II in Tabelle 2). Nach Berücksichtigung der Einflüsse der Variablen des familiären Hintergrunds (vgl. wiederum Tabelle 1) wird der Zusammenhang zwischen Computern und PISA-Leistungen sehr klein und statistisch insignifikant (Spalte III in Tabelle 2). Und wenn schließlich auch noch die Einflüsse häuslicher und schulischer Inputs sowie institutioneller Gegebenheiten berücksichtigt werden, dann kehrt sich der Zusammenhang gar in den in Tabelle 1 berichteten statistisch signifikant negativen Effekt um.

Dieses Ergebnismuster zeigt auf, wie sehr einfache bivariate Betrachtungen irreführend sein können – auch wenn sie selbst in hochqualitativen Dokumenten wie den offiziellen PISA-Veröffentlichungen berichtet werden.¹⁵ Im vorliegenden Fall ist es offensichtlich, dass die Verfügbarkeit von Computern zu Hause stark vom Bildungsniveau und Beruf der Eltern abhängt. Gut ausgebildete Eltern legen aber auch mehr Wert auf die Ausbildung ihrer Kinder. Daher ist es nicht verwunderlich, dass Kinder gut ausgebildeter Eltern sowohl mehr Zugang zu einem Computer haben als auch bessere Schulleistungen erbringen als Kinder aus bildungsferneren Schichten. Um den Einfluss der Computerverfügbarkeit auf die Schulleistungen zu bestimmen, muss man die Leistungen von Schülern aus ähnlichen familiären Verhältnissen miteinander vergleichen, was in der multivariaten Regressionsanalyse geschieht. Und dabei ergibt sich dann eben kein positiver, sondern sogar ein negativer Zusammenhang.

¹⁴ Das Ergebnismuster in Naturwissenschaften und Lesen entspricht ganz dem in Mathematik.

¹⁵ Vgl. OECD (2001, S. 118), die – trotz Vorsichtsäußerungen bezüglich möglicher Interpretationsprobleme bivariater Zusammenhänge – folgenden bivariaten Befund berichten: „Students with higher values on the index of interest in computers tend to perform better on the combined reading literacy scale.“

Dies deutet darauf hin, dass gerade zu Hause Computer weniger für Lernzwecke als vielmehr zum Spielen von Computerspielen oder ähnlichen Freizeitbeschäftigungen genutzt werden (vgl. dazu Wirth und Klieme 2003). Wenn dies dann auf Kosten der Hausarbeiten oder sonstiger lernfördernder Betätigungen geht, ist der negative Zusammenhang zwischen Computerverfügbarkeit zu Hause und den PISA-Leistungen in den drei Kompetenzbereichen durchaus verständlich. Dies legt auch eine differenziertere Analyse des Zusammenhangs zwischen PISA-Leistungen und der Nutzung des häuslichen Computers nahe (vgl. Fuchs und Wößmann 2004b), welche positive Zusammenhänge mit der Nutzung von Bildungssoftware sowie mit der Nutzung des Computers für Kommunikations- und Informationszwecke per Email und Internet aufzeigt.

Neben weiteren Details berichten Fuchs und Wößmann (2004b) auch Ergebnisse zum Zusammenhang zwischen PISA-Leistungen und der Verfügbarkeit und Nutzung von Computern in der Schule. Dabei ergibt sich bei multivariater Betrachtung kein signifikanter Zusammenhang zur Verfügbarkeit von Computern in der Schule. In Bezug auf die Nutzungsintensität von Computern und Internet in der Schule ergibt sich ein Zusammenhang in Form eines umgekehrten „U“: Die PISA-Leistungen nehmen leicht zu, wenn man von einer Intensität von so gut wie keiner Computernutzung zu einer häufigeren Intensität von bis zu mehrmals im Monat geht, doch sie nehmen stark ab, wenn man eine Intensität der Computernutzung von mehrmals pro Woche betrachtet. Eine Erklärung hierfür könnte sein, dass die starke Nutzung des Computers im Unterricht andere effektivere Lernaktivitäten verdrängt.

3.3 Schulische Ressourcen und Lehrereigenschaften

Es besteht ein positiver Zusammenhang zwischen den PISA-Leistungen in Mathematik und Naturwissenschaften und dem durchschnittlichen Niveau der Bildungsausgaben eines Landes (vgl. Tabelle 1). Allerdings ist dieser Zusammenhang quantitativ recht gering (10,7 Punkte in Mathematik und 7,7 Punkte in Naturwissenschaft für 1.000 Dollar zusätzliche jährliche Ausgaben) und wird hauptsächlich von einigen Ländern mit außergewöhnlich geringen Ausgaben getrieben, besteht aber nicht zwischen den entwickelten OECD-Ländern und auch nicht im Lesetest (vgl. Fuchs und Wößmann 2004a für Details).

In Bezug auf die Klassengröße findet sich kein negativer Zusammenhang zwischen Klassengröße und PISA-Leistungen (schlechtere Leistungen in größeren Klassen), sondern in Mathematik und Naturwissenschaften sogar genau umgekehrt ein signifikant positiver Zusammenhang (bessere Leistungen in größeren Klassen). Dies dürfte allerdings vor allem dar-

auf zurückzuführen sein, dass vielerorts eine systematische Zuordnung schwächerer Schüler in Schulen mit kleineren Klassen stattfindet (vgl. Abschnitt 2.1 sowie West und Wößmann 2005).

Hingegen sind die PISA-Leistungen dort besser, wo geeignete Unterrichtsmaterialien (wie etwa Lehrbücher) nicht fehlen. In Mathematik und Naturwissenschaften ergibt sich auch ein positiver Zusammenhang mit der jährlichen Unterrichtszeit.¹⁶ Darüber hinaus besteht ein signifikant positiver Zusammenhang der PISA-Leistungen in allen drei Kompetenzbereichen mit dem Bildungsstand der Lehrer, wobei der Effekt eines Universitätsabschlusses in Erziehungswissenschaften zumeist geringer ist als der von spezifischen Lehrertiteln und von Universitätsabschlüssen im jeweiligen Unterrichtsfach.

Dementsprechend scheint ein positiver Zusammenhang zwischen Schülerleistungen und schulischer Ressourcenausstattung weitgehend auf die Qualität der vorhandenen Ressourcen (Lehrer und Lehrmaterialien) beschränkt zu sein. Der fehlende eindeutige Zusammenhang zwischen PISA-Leistungen und quantitativen Ressourcenmaßen wie Ausgaben pro Schüler oder Klassengrößen im Länderquerschnitt spiegelt hingegen bisherige Befunde wider, die ebenfalls keinen Zusammenhang zwischen Schülerleistungen und Ausgabenniveaus über die Zeit feststellen konnten (siehe Gundlach und Wößmann 2000; Gundlach et al. 2001).

3.4 Institutionelle Effekte

Wie schon in der Einleitung angeführt, legt die ökonomische Theorie nahe, dass institutionelle Gegebenheiten des Schulsystems aufgrund ihrer Auswirkungen auf die Anreize der am Bildungsprozess Beteiligten die erzielten Schülerleistungen beeinflussen können. Dementsprechend zeigen die empirischen Befunde einen signifikant positiven Zusammenhang zwischen der Existenz von zentralen Abschlussprüfungssystemen in einem Land und den Schülerleistungen in Mathematik und Naturwissenschaften.¹⁷ Während zentrale Abschlussprüfungssysteme zumeist nationale Merkmale sind,¹⁸ liegen im PISA-Datensatz auch Informatio-

¹⁶ Der negative Zusammenhang der in PISA gemessenen Lesefähigkeiten mit den Sprach-Unterrichtsstunden könnte damit zusammenhängen, dass Sprachunterricht nicht unbedingt die in PISA umfassend definierte „Lesekompetenz“ lehrt.

¹⁷ Der Zusammenhang in Naturwissenschaften ist statistisch signifikant auf dem 11%-Niveau. Die Tatsache, dass der positive Zusammenhang mit den Leseleistungen nicht statistisch signifikant ist, dürfte stark damit zusammenhängen, dass keine international vergleichbaren Daten über die Existenz zentraler Abschlussprüfungen in sprachlichen Fächern vorlagen und statt dessen Bishop (2005) folgend nur ein einfacher Durchschnitt des Vorliegens zentraler Prüfungen in Mathematik und Naturwissenschaften verwendet wurde.

¹⁸ Vgl. aber für Deutschland Büchel et al. (2005) und Jürges et al. (2004), die ähnliche Befunde positiver Effekte von Zentralprüfungen zwischen den deutschen Bundesländern anhand der deutschen TIMSS- und PISA-E-Daten finden.

nen auf Schulebene darüber vor, ob 15jährige Schüler zumindest einmal jährlich einem standardisierten Test unterzogen wurden. Bei Kontrolle für zentrale Abschlussprüfungen ist der Zusammenhang zwischen regelmäßigen standardisierten Tests und den PISA-Leistungen statistisch nicht signifikant. Betrachtet man die regelmäßigen standardisierten Tests aber als alternatives Maß für externe Überprüfung und lässt die Variable zu zentralen Abschlussprüfungen aus der Schätzgleichung, dann ergibt sich ein statistisch signifikanter positiver Zusammenhang mit den Schülerleistungen in allen drei PISA-Kompetenzbereichen. Die Ergebnisse legen also nahe, dass zentrale Prüfungssysteme mit besseren Schülerleistungen einhergehen.¹⁹

In Bezug auf die Autonomie von Schulen ergeben sich in verschiedenen Entscheidungsbereichen sehr unterschiedliche Ergebnisse.²⁰ So zeigt sich im Bereich von Prozess- und Personalentscheidungen zumeist ein positiver Zusammenhang zwischen PISA-Leistungen und Schulautonomie. Dies ist zum Beispiel bei der Auswahl der Lehrbücher (mit quantitativ sehr großen Effekten), der Bestimmung der Budgetverteilung innerhalb der Schule und bei der Einstellung von Lehrern der Fall.²¹ Demgegenüber besteht ein negativer Zusammenhang zwischen PISA-Leistungen und Schulautonomie in der Formulierung des Schulbudgets. Diese Kombination von Ergebnissen legt nahe, dass es leistungsfördernd ist, wenn den Schulen der Budgetrahmen vorgegeben ist, sie innerhalb des Rahmens aber selbständig entscheiden können. Ganz ähnliche Befunde konnten auch schon für die TIMSS-Ergebnisse berichtet werden (Wößmann 2003a). Obwohl es in der hier betrachteten einfachen Modellspezifikation zumeist keinen signifikanten Zusammenhang zwischen Schülerleistungen und Schulautonomie in Gehaltsfragen sowie in der Bestimmung von Lehrinhalten gibt, so verdeckt dies doch wichtige signifikante Interaktionseffekte in diesen Bereichen, die im folgenden Abschnitt angesprochen werden.

Schließlich liefert der PISA-Datensatz Informationen darüber, ob die jeweiligen Schulen in öffentlicher oder privater Trägerschaft operieren sowie wie hoch ihr öffentlicher und privater Finanzierungsanteil sind. Im Durchschnitt schneiden Schüler an Schulen in privater Trägerschaft in allen drei PISA-Kompetenzbereichen statistisch signifikant besser ab als Schüler an

¹⁹ Ergebnisse von Quantilregressionen legen darüber hinaus nahe, dass der Effekt zentraler Abschlussprüfungen für gute Schüler etwas größer ist als für schwache Schüler (Wößmann 2005a).

²⁰ Vgl. Bishop und Wößmann (2004) für eine theoretische Erklärung dieser variierenden Ergebnisse zur Schulautonomie.

²¹ Wird der Effekt von Schulautonomie in der Einstellung von Lehrern herausgerechnet, so ergibt sich kein signifikanter Zusammenhang zwischen Schulautonomie in der Entlassung von Lehrern und den PISA-Leistungen in Mathematik und Naturwissenschaften, und sogar ein leicht negativer in Lesen. Hinsichtlich der Problematik der Interpretation dieser Variablen und für zusätzliche Ergebnisse vor allem in Bezug auf Effektheterogenität vgl. Fuchs und Wößmann (2004a).

Schulen in öffentlicher Trägerschaft.²² Aber sobald diese Effekte der Form der Trägerschaft berücksichtigt werden, besteht kein signifikanter Zusammenhang zwischen dem Anteil der öffentlichen Finanzierung und den Schülerleistungen in Naturwissenschaften und Lesen. In Mathematik ist es sogar so, dass Schüler in Schulen mit einem höheren Anteil öffentlicher Finanzierung besser abschneiden.

3.5 Die Heterogenität des Effekts von externen Abschlussprüfungen

Das bisher betrachtete Modell erlaubte keine Heterogenität der Effekte zwischen verschiedenen Schulsystemen. Diese Restriktion soll nun ein Stück weit gelockert werden, indem systematische Unterschiede in den betrachteten Zusammenhängen zwischen Schulsystemen mit zentralen Abschlussprüfungssystemen und solchen ohne zentrale Abschlussprüfungen zugelassen werden. Dazu wird Gleichung (3) aus Abschnitt 2.1 geschätzt, die Interaktionseffekte zwischen externen Abschlussprüfungen und den anderen Einflussfaktoren zulässt. Details zu der Spezifikation und zusätzliche Ergebnisse zum Interaktionsmodell finden sich in Wößmann (2005a). Hier soll beispielhaft auf besonders interessante Interaktionen zwischen zentralen Prüfungen und anderen institutionellen Gegebenheiten des Schulsystems, nämlich Schulautonomie und regelmäßige standardisierte Tests, eingegangen werden.

Aus theoretischer Sicht können solche Effektheterogenitäten deshalb erwartet werden, weil externe Prüfungen die im Bildungssystem bestehenden Informationsasymmetrien mildern können und mit dieser Transparenz ein gewisses Maß an Rechenschaft und Verantwortlichkeit („accountability“) in das System einführen, welches allzu opportunistisches Verhalten dezentraler Entscheidungsträger verhindert. Damit kann es zu Komplementaritäten zwischen externen Prüfungen und Schulautonomie kommen, deren Ausmaß durch die Anreize für lokales opportunistisches Verhalten und das Vorliegen lokaler Wissensvorsprünge im jeweiligen Entscheidungsbereich bestimmt werden (vgl. Wößmann 2003c; 2005a).

Die empirischen Schätzergebnisse werden in den Abbildungen 1-4 beispielhaft graphisch dargestellt (s. Wößmann 2005a für detaillierte Ergebnisse). Wie Abbildung 1 zeigt, besteht in Systemen ohne zentrale Prüfungen ein negativer Zusammenhang zwischen Schülerleistungen und Schulautonomie über die Lehrergehälter. In Systemen mit zentralen Prüfungen dreht sich dieser Effekt komplett um: In Zentralprüfungssystemen sind die Schülerleistungen besser in

²² Dieser Befund erweist sich als robust in einer Schätzung, die zusätzlich für die Zusammensetzung der Schülerpopulation (in Form des Medians der elterlichen Bildung und des sozioökonomischen Status auf Schulebene) kontrolliert. Dies stützt den Befund von Dronkers und Roberts (2003), die ebenfalls darlegen, dass der Leistungsunterschied zwischen öffentlichen und privaten Schulen nicht aufgrund der unterschiedlichen Schülerzusammensetzung zustande kommt.

Schulen, die Gehaltsautonomie haben. In Systemen ohne zentrale Prüfungen sind die PISA-Leistungen ebenfalls geringer in Schulen, die Eigenverantwortung über die Lehrinhalte haben (Abbildung 2). Wiederum dreht sich dieser Zusammenhang in einen positiven um, wenn zentrale Abschlussprüfungen vorliegen. Wie Abbildung 3 belegt, besteht zwar in beiden Prüfungssystemen ein positiver Zusammenhang zwischen Schülerleistungen und Schulautonomie über Budgetverteilung innerhalb der Schule. Aber dieser Zusammenhang ist wesentlich stärker und statistisch signifikant in Systemen mit zentralen Abschlussprüfungen. Der Unterschied in den Auswirkungen der Schulautonomie zwischen den beiden Prüfungssystemen ist in allen drei Fällen statistisch signifikant.

Aus ökonomischer Sicht lässt sich der diametral unterschiedliche Einfluss von Schulautonomie in Systemen mit und ohne zentralen Abschlussprüfungen dadurch erklären, dass die Anreizbedingungen für Schulen zwischen den beiden Systemen differieren. In Systemen ohne zentrale Prüfungen kann opportunistisches Verhalten der Schulen für gewöhnlich nicht festgestellt werden und wird daher nicht sanktioniert. Demgegenüber schaffen zentrale Prüfungen Informationen darüber, ob die Schulen gute Leistungsergebnisse erzielen oder nicht, so dass Aufsichtsbehörden und gegebenenfalls Eltern mögliche Konsequenzen aus leistungsminderndem Verhalten der Schulen ziehen können. Dies schafft Anreize für die Entscheidungsträger in den Schulen, ihre Autonomie nicht opportunistisch auszunutzen, sondern effektiv zur Leistungsförderung der Schüler einzusetzen. Somit dürften die unterschiedlichen Effekte der Schulautonomie in Systemen mit und ohne zentralen Abschlussprüfungen Verhaltensänderungen der schulischen Entscheidungsträger widerspiegeln.

In diesem Sinne können zentrale Abschlussprüfungen als „Währung“ des Schulsystems fungieren, die sich als zentraler Wertmaßstab geradezu als Voraussetzung dafür erweisen, dass dezentral organisierte Schulsysteme zu hohen Schülerleistungen führen können. Eine effiziente Bildungspolitik sollte also Zentralprüfungen mit Schulautonomie verbinden: Sie sollte Standards extern vorgeben und überprüfen und es gleichzeitig den Schulen überlassen, wie sie diese Standards erreichen wollen.

In Abbildung 4 zeigt sich schließlich ein ähnliches Bild für den Fall des Einsatzes regelmäßiger standardisierter Tests bei den 15jährigen Schülern. Während in Systemen ohne zentrale Prüfungen ein negativer Zusammenhang zwischen den getesteten Schülerleistungen und standardisierten Tests bestehen, ist dieser in Systemen mit zentralen Prüfungen signifikant positiv. Dies könnte darauf hin deuten, dass solche regelmäßigen standardisierten Tests nur

dann leistungsfördernd sind, wenn durch die zentralen Abschlussprüfungssysteme die Ziele und Standards des Schulsystems klar vorgegeben sind.

3.6 Erklärungsgehalt zwischen den Teilnehmerstaaten

Da die Fragestellung des vorliegenden Beitrags sich auf die Bestimmungsgründe der internationalen Unterschiede in Schülerleistungen bezieht, stellt sich abschließend die Frage, welchen Anteil der internationalen Leistungsstreuung das geschätzte Modell denn erklären kann. Fuchs und Wößmann (2004a) zeigen dazu eine Methode, die den Erklärungsgehalt des Modells zwischen den Teilnehmerstaaten (im Unterschied zu zwischen den individuellen Schülern) abbilden kann. Dabei ergibt sich, dass das vorgestellte Modell 88 Prozent der zwischenstaatlichen Leistungsstreuung in Mathematik, 86 Prozent in Naturwissenschaften und 85 Prozent in Lesen erklären kann (R^2 des Modells auf Staatsebene). Somit machen die berücksichtigten Ursachenkomplexe – familiärer Hintergrund, schulische Ressourcenausstattung und institutionelle Ausgestaltung – in der Tat fast die gesamte Leistungsvariation zwischen den PISA-Teilnehmerstaaten aus.

Dabei tragen alle betrachteten Ursachenkomplexe signifikant zu der Erklärung bei. Je nach Modell und Betrachtungsweise lassen sich rund ein Viertel der gesamten internationalen Streuung auf internationale Unterschiede in den betrachteten institutionellen Rahmenbedingungen der Schulsysteme zurückführen. Insbesondere in der Lesefähigkeit ist dies wesentlich bedeutsamer als etwa der Erklärungsbeitrag der internationalen Unterschiede in der Ressourcenausstattung der Schulen, wobei sogar noch zu berücksichtigen ist, dass Letzterer teilweise durch die erwartungskonforme positive Korrelation zwischen Leistungen und Klassengrößen bedingt ist (vgl. Fuchs und Wößmann 2004a für weitere Details).

4. Schlussbemerkungen

Der vorliegende Beitrag berichtete die Ergebnisse umfangreicher Untersuchungen der internationalen PISA-Mikrodaten in Bezug auf mögliche Ursachenkomplexe der Leistungsunterschiede. Das vorgestellte Modell, das zahlreiche Einflussfaktoren in den drei Bereichen familiärer Hintergrund der Schüler, Ressourcenausstattung der Schulen und institutionelle Gestaltung des Schulsystems berücksichtigt, kann den Großteil der internationalen Unterschiede in den PISA-Leistungen in Mathematik, Naturwissenschaften und Lesen begründen. Dabei sind die Einflüsse institutioneller Rahmenbedingungen von besonderem Interesse. Rund ein Viertel der gesamten Leistungsunterschiede zwischen den PISA-Teilnehmerländern lassen

sich auf internationale Unterschiede in Prüfungssystemen, Schulautonomie und Schulträgerschaft zurückführen. Insbesondere schneiden Schüler in Systemen mit zentralen Abschlussprüfungen und in Schulen in privater Trägerschaft besser ab. Darüber hinaus besteht auch ein positiver Zusammenhang zwischen den PISA-Leistungen und Schulautonomie in Personal- und Prozessentscheidungen, insbesondere dort, wo durch zentrale Prüfungssysteme eine Rechenschaftsfunktion gegeben ist.

Natürlich kann eine solche vereinheitlichende Betrachtungsweise nicht die ganze Komplexität des Bildungsgeflechts widerspiegeln, und aufgrund der einfachen Querschnittsstruktur der PISA-Daten lassen sich die Ergebnisse nur als differenzierte deskriptive Zusammenhänge interpretieren, die empirisch gesehen nicht notwendigerweise direkte kausale Beziehungen darstellen. Aber die präsentierten Ergebnisse liefern erste Ideen über wichtige durchgreifende Muster von Zusammenhängen, die sich erkenntnisleitend einsetzen lassen. Für ein umfassenderes Bild sind selbstverständlich Detailanalysen innerhalb der verschiedenen Systeme unerlässlich. Diesbezüglich sind die eingehenden Arbeiten zu verschiedenen Ländern in Ammermüller et al. (2005), Wößmann (2005b; 2005c), Wößmann und Fuchs (2005) und Wößmann und West (2005) von Interesse.

Auch wurden in den vorliegenden Untersuchungen aufgrund der gegebenen Datenbasis sicherlich nicht alle relevanten institutionellen Gegebenheiten der Schulsysteme analysiert. Neben den hier betrachteten Institutionen ließen sich etwa auch die Einflüsse von schulbezogenen Rechenschaftssystemen oder konkreten Anreizsystemen für Lehrer untersuchen (vgl. Wößmann 2004a für einen Überblick).

Schließlich lag der Fokus der vorgestellten Untersuchungen insbesondere auf bildungspolitischen Maßnahmen und dabei auf Fragen der effizienten Bildungsbereitstellungen. Eine zusätzliche Analyse möglicher Einflüsse familienpolitischer Maßnahmen wäre von großem Interesse, ebenso wie Fortschritte bei der Identifikation pädagogisch bedeutsamer Einflussfaktoren wie etwa Lehrinhalte und -methoden. Für Untersuchungen der internationalen Unterschiede in der Chancengleichheit im Schulbereich und ihrer möglichen Bestimmungsgründe sei auf Wößmann (2004b), Hanushek und Wößmann (2005), Schütz und Wößmann (2005) sowie Peterson und Wößmann (2005) verwiesen. In diesen Beiträgen werden unter anderem auch Einflüsse weiterer institutioneller Gegebenheiten wie des Umfangs des Systems frühkindlicher Bildung und der Mehrgliedrigkeit des Schulsystems analysiert.

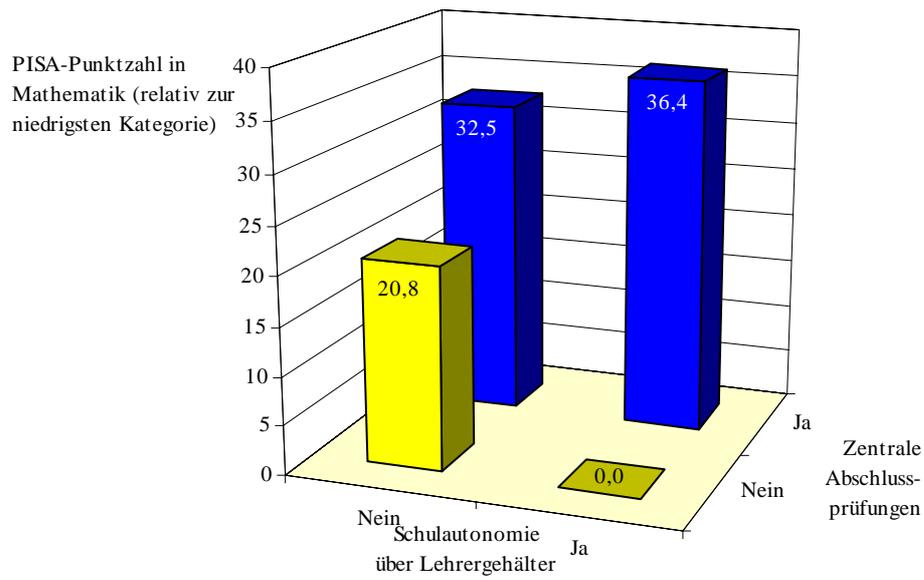
Literatur

- Adams, Ray, Margaret Wu (eds.) (2002). *PISA 2000 Technical Report*. Paris: Organisation for Economic Co-operation and Development (OECD).
- Akerhielm, Karen (1995). Does Class Size Matter? *Economics of Education Review* 14 (3): 229-241.
- Ammermüller, Andreas, Hans Heijke, Ludger Wößmann (2005). Schooling Quality in Eastern Europe: Educational Production During Transition. *Economics of Education Review* 24 (5): 579-599.
- Barro, Robert J. (2001). Human Capital and Growth. *American Economic Review, Papers and Proceedings* 91 (2): 12-17.
- Bishop, John H. (1997). The Effect of National Standards and Curriculum-Based Exams on Achievement. *American Economic Review* 87 (2): 260-264.
- Bishop, John H. (2005). Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn. Erscheint in: Eric A. Hanushek, Finis Welch (Hrsg.), *Handbook of the Economics of Education*. Amsterdam: North-Holland.
- Bishop, John H., Ludger Wößmann (2004). Institutional Effects in a Simple Model of Educational Production. *Education Economics* 12 (1): 17-38.
- Bryk, Anthony S., Stephen W. Raudenbush (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Advanced Quantitative Techniques in the Social Sciences 1. Newbury Park: Sage Publications.
- Büchel, Felix, Hendrik Jürges, Kerstin Schneider (2005). The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany. Erscheint in: *Journal of the European Economic Association*. (Working-Paper-Version verfügbar als CESifo Working Paper 939 unter www.cesifo.de)
- Cohen, Jon, Stéphane Baldi (1998). An Evaluation of the Relative Merits of HLM vs. Robust Linear Regression in Estimating Models with Multi-Level Data. Mimeo, American Institutes for Research.
- Deaton, Angus (1997). *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore: The Johns Hopkins University Press.
- Dronkers, Jaap, Péter Robert (2003). The Effectiveness of Public and Private Schools from a Comparative Perspective. EUI Working Paper SPS 2003-13. Florence: European University Institute.
- DuMouchel, William H., Greg J. Duncan (1983). Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples. *Journal of the American Statistical Association* 78 (383): 535-543.
- Entorf, Horst, Nicoleta Minoiu (2005). What a Difference Immigration Policy Makes: A Comparison of PISA Scores in Europe and Traditional Countries of Immigration. Erscheint in: *German Economic Review*. (Working-Paper-Version verfügbar als IZA Discussion Paper 1021 unter www.iza.org)
- Fertig, Michael, Robert E. Wright (2005). School Quality, Educational Attainment and Aggregation Bias. Erscheint in: *Economics Letters*. (Working-Paper-Version verfügbar als IZA Discussion Paper 994 unter www.iza.org)
- Fuchs, Thomas, Ludger Wößmann (2004a). What Accounts for International Differences in Student Performance? A Re-examination using PISA Data. CESifo Working Paper 1235. Munich: CESifo. (Verfügbar unter www.cesifo.de)
- Fuchs, Thomas, Ludger Wößmann (2004b). Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and Use of Computers at Home and at School. Erscheint in: *Brussels Economic Review* 47 (3). (Working-Paper-Version verfügbar als CESifo Working Paper 1321 unter www.cesifo.de)

- Gundlach, Erich, Ludger Wößmann (2000). Die sinkende Produktivität der Schulen in OECD-Ländern. In: Manfred Weiß, Horst Weishaupt (Hrsg.), *Bildungsökonomie und Neue Steuerung*, Beiträge zur Bildungsplanung und Bildungsökonomie Vol. 9, S. 31-52. Frankfurt: Lang.
- Gundlach, Erich, Ludger Wößmann (2004). Bildungsressourcen, Bildungsinstitutionen und Bildungsqualität: Makroökonomische Relevanz und mikroökonomische Evidenz. In: Uschi Backes-Gellner, Petra Moog (Hrsg.), *Ökonomie der Evaluation von Schulen und Hochschulen*. Schriften des Vereins für Socialpolitik, N.F. Band 302, S. 15-52. Berlin: Duncker & Humblot.
- 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 (Hrsg.), *Handbook of Public Economics, Volume 4*: 2045-2141. Amsterdam: North Holland.
- Hanushek, Eric A., Dennis D. Kimko (2000). Schooling, Labor-Force Quality, and the Growth of Nations. *American Economic Review* 90 (5): 1184-1208.
- Hanushek, Eric A., Ludger Wößmann (2005). Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries. NBER Working Paper 11124. Cambridge, Mass.: National Bureau of Economic Research. (Verfügbar unter www.nber.org)
- Jürges, Hendrik, Wolfram F. Richter, Kerstin Schneider (2004). Teacher Quality and Incentives: Theoretical and Empirical Effects of Standards on Teacher Quality. CESifo Working Paper 1296. Munich: CESifo. (Verfügbar unter www.cesifo.de)
- Lee, Jong-Wha, Robert J. Barro (2001). Schooling Quality in a Cross-Section of Countries. *Economica* 68 (272): 465-488.
- Moulton, Brent R. (1986). Random Group Effects and the Precision of Regression Estimates. *Journal of Econometrics* 32 (3): 385-397.
- Organisation for Economic Co-operation and Development (OECD) (2000). *Measuring Student Knowledge and Skills: The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy*. Paris: OECD.
- Organisation for Economic Co-operation and Development (OECD) (2001). *Knowledge and Skills for Life: First Results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris: OECD.
- Organisation for Economic Co-operation and Development (OECD) (2002). *Manual for the PISA 2000 Database*. Paris: OECD.
- Peterson, Paul, Ludger Wößmann (Hrsg.) (2005). *Schools and the Equal Opportunity Problem*. Erscheint bei: Cambridge, MA: MIT Press.
- Schütz, Gabriela, Ludger Wößmann (2005). Chancengleichheit im Schulsystem: Internationale deskriptive Evidenz und mögliche Bestimmungsfaktoren. Ifo Working Paper (im Druck). München: ifo Institut für Wirtschaftsforschung an der Universität München. (In Kürze verfügbar unter www.cesifo.de)
- Stock, James H., Mark W. Watson (2003). *Introduction to Econometrics*. Boston: Pearson Education/Addison-Wesley.
- West, Martin R., Ludger Wößmann (2005). Which School Systems Sort Weaker Students into Smaller Classes? International Evidence. Erscheint in: *European Journal of Political Economy*. (Working-Paper-Version verfügbar als CESifo Working Paper 1054 unter www.cesifo.de)
- White, Halbert (1984). *Asymptotic Theory for Econometricians*. Orlando: Academic Press.

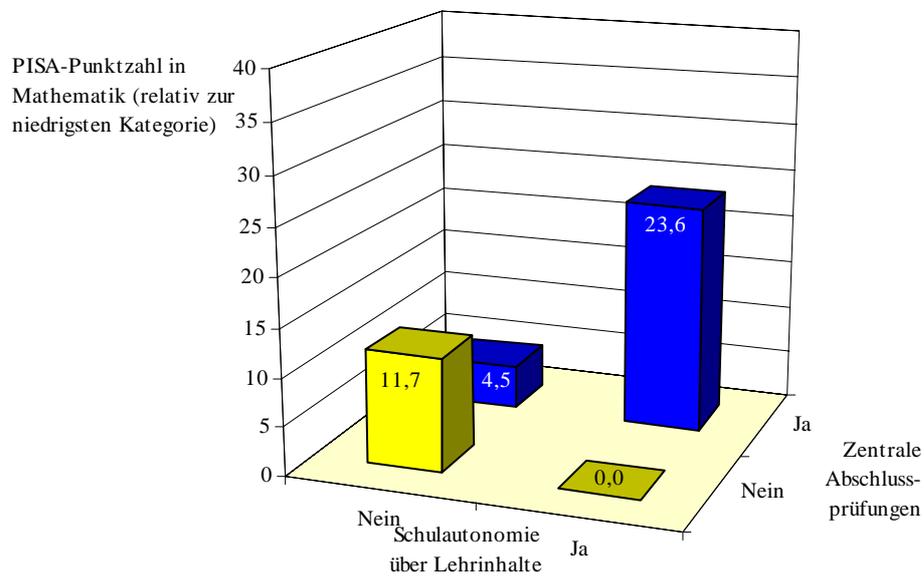
- Wirth, Joachim, Eckhard Klieme (2003). Computernutzung. In: Deutsches PISA-Konsortium (Hrsg.), *PISA 2000: Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*: 195-209. Opladen: Leske + Budrich.
- Wolter, Stefan C., Maja Coradi Vellacott (2003). Sibling Rivalry for Parental Resources: A Problem for Equity in Education? A Six-Country Comparison with PISA Data. *Swiss Journal of Sociology* 29 (3): 377-398.
- Wooldridge, Jeffrey M. (2001). Asymptotic Properties of Weighted M -Estimators for Standard Stratified Samples. *Econometric Theory* 17 (2): 451-470.
- Wößmann, Ludger (2001a). Why Students in Some Countries Do Better: International Evidence on the Importance of Education Policy. *Education Matters* 1 (2): 67-74.
- Wößmann, Ludger (2001b). Schulsystem und Schülerleistung im internationalen Vergleich: Was Institutionen ausmachen. *Die Weltwirtschaft* (3): 283-304.
- Wößmann, Ludger (2002). *Schooling and the Quality of Human Capital*. Berlin: Springer.
- Wößmann, Ludger (2003a). Schooling Resources, Educational Institutions and Student Performance: the International Evidence. *Oxford Bulletin of Economics and Statistics* 65 (2): 117-170.
- Wößmann, Ludger (2003b). Central Exit Exams and Student Achievement: International Evidence. In: Paul E. Peterson, Martin R. West (eds.), *No Child Left Behind? The Politics and Practice of School Accountability*. Washington, D.C.: Brookings Institution Press.
- Wößmann, Ludger (2003c). Zentrale Prüfungen als „Währung“ des Bildungssystems: Zur Komplementarität von Schulautonomie und Zentralprüfungen. *Vierteljahreshefte zur Wirtschaftsforschung* 72 (2): 220-237.
- Wößmann, Ludger (2003d). Specifying Human Capital. *Journal of Economic Surveys* 17 (3): 239-270.
- Wößmann, Ludger (2003e). Familiärer Hintergrund, Schulsystem und Schülerleistungen im internationalen Vergleich. *Aus Politik und Zeitgeschichte* (21/22): 33-38.
- Wößmann, Ludger (2004a). Institutional Comparisons in Educational Production. *CESifo DICE Report - Journal for Institutional Comparisons* 2 (4): 3-6.
- Wößmann, Ludger (2004b). How Equal Are Educational Opportunities? Family Background and Student Achievement in Europe and the United States. CESifo Working Paper 1162. Munich: CESifo. (Verfügbar unter www.cesifo.de)
- Wößmann, Ludger (2005a). The Effect Heterogeneity of Central Exams: Evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics* 13 (2): 143-169.
- Wößmann, Ludger (2005b). Educational Production in East Asia: The Impact of Family Background and Schooling Policies on Student Performance. Erscheint in: *German Economic Review*. (Working-Paper-Version verfügbar als ICSEAD Working Paper 2003-17 unter www.icsead.or.jp)
- Wößmann, Ludger (2005c). Educational Production in Europe. *Economic Policy* 20 (43): 445-504.
- Wößmann, Ludger, unter Mithilfe von Thomas Fuchs (2005). Families, Schools, and Primary-School Learning: Evidence for Argentina and Colombia in an International Perspective. World Bank Policy Research Working Paper 3537. Washington, D.C.: World Bank. (Verfügbar unter www-wds.worldbank.org)
- Wößmann, Ludger, Martin R. West (2005). Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS. Erscheint in: *European Economic Review*.

Abbildung 1: Zentrale Abschlussprüfungen und Gehaltsautonomie der Schulen



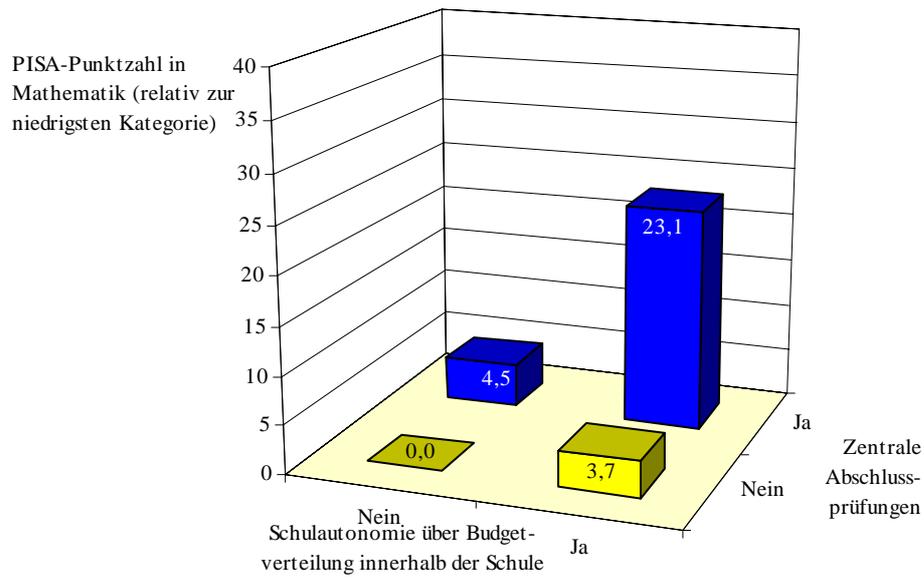
Quelle: Wößmann (2005a)

Abbildung 2: Zentrale Abschlussprüfungen und Schulautonomie über Lehrinhalte



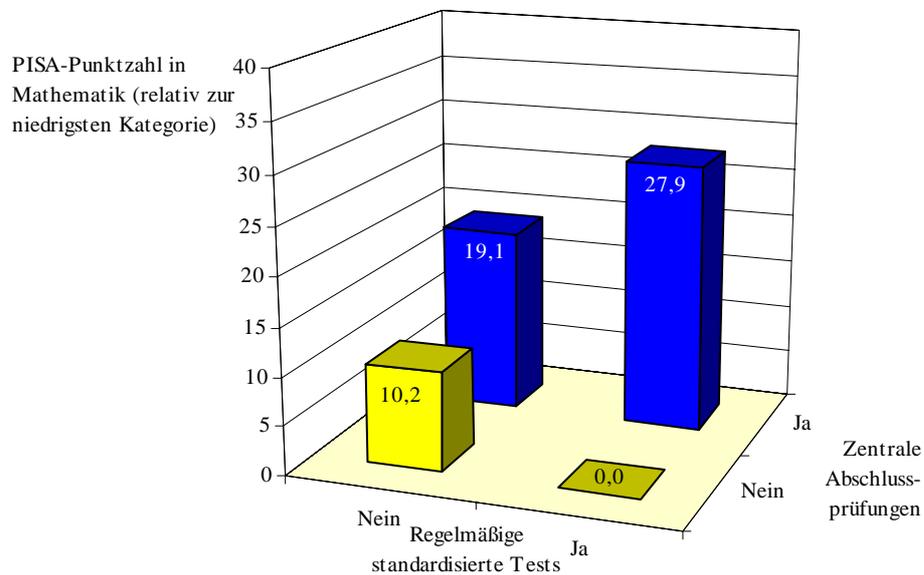
Quelle: Wößmann (2005a)

Abbildung 3: Zentrale Abschlussprüfungen und Budgetverteilung innerhalb der Schulen



Quelle: Eigene Berechnungen anhand der PISA-Daten.

Abbildung 4: Zentrale Abschlussprüfungen und regelmäßige standardisierte Tests



Quelle: Wößmann (2005a)

Tabelle 1: Internationale „Bildungsproduktionsfunktionen“

	Mathematik		Naturwissenschaften		Lesen	
	Koeffizient	Std.fehler	Koeffizient	Std.fehler	Koeffizient	Std.fehler
Schülereigenschaften						
<i>Klasse (ausgelassen: 10. Klasse)</i>						
6. oder niedriger	-99,4 ***	(12,5)	-86,5 ***	(11,2)	-175,8 ***	(22,7)
7.	-97,6 ***	(5,6)	-73,2 ***	(4,2)	-106,2 ***	(5,0)
8.	-69,6 ***	(2,7)	-60,1 ***	(2,3)	-85,7 ***	(3,9)
9.	-25,3 ***	(1,7)	-23,9 ***	(1,6)	-34,5 ***	(2,3)
11.	10,3 ***	(2,2)	11,1 ***	(2,4)	8,4 ***	(2,2)
12. oder höher	68,0 ***	(6,6)	71,3 ***	(8,5)	56,0 ***	(6,3)
Alter (Monate)	-0,4 ***	(0,1)	0,1	(0,1)	-0,4 ***	(0,1)
Weiblich	-16,1 ***	(0,9)	-4,0 ***	(0,8)	23,9 ***	(0,8)
Familiärer Hintergrund						
<i>Im Land geboren</i>						
Schüler	4,8 **	(2,1)	7,0 ***	(2,0)	12,4 ***	(2,1)
Mutter	6,5 ***	(1,5)	8,1 ***	(1,6)	8,9 ***	(1,3)
Vater	5,6 ***	(1,5)	11,7 ***	(1,5)	8,2 ***	(1,3)
<i>Schüler lebt mit (ausgelassen: keinem Elternteil)</i>						
Alleinerziehendem Vater	19,0 ***	(3,7)	19,0 ***	(3,5)	26,1 ***	(4,7)
Alleinerziehender Mutter	6,7	(4,2)	12,4 ***	(4,2)	14,0 ***	(4,6)
Beiden Elternteilen	14,0 ***	(3,7)	17,4 ***	(3,5)	22,2 ***	(4,4)
<i>Bildungsstand der Eltern (ausgelassen: keine formale Bildung)</i>						
Primar	11,9 ***	(4,0)	11,8 ***	(4,1)	18,4 ***	(3,9)
Untere Sekundar	15,6 ***	(4,1)	12,9 ***	(4,0)	20,4 ***	(3,9)
Höhere Sekundar 1	17,8 ***	(4,2)	16,9 ***	(4,1)	26,9 ***	(3,9)
Höhere Sekundar 2	21,8 ***	(4,1)	19,7 ***	(4,1)	28,1 ***	(3,9)
Universität	26,9 ***	(4,2)	26,5 ***	(4,1)	34,3 ***	(3,9)
<i>Arbeitsstatus der Eltern (ausgelassen: arbeitslos)</i>						
Mindestens einer halbtags	0,5	(2,0)	-3,1	(2,0)	-1,8	(1,8)
Mindestens einer ganztags	15,8 ***	(1,7)	10,6 ***	(1,6)	11,7 ***	(1,5)
Beide ganztags	14,0 ***	(1,7)	11,0 ***	(1,7)	10,8 ***	(1,5)
<i>Beruf der Eltern (ausgelassen: Zwischengröße)</i>						
„Blue collar“	-12,4 ***	(1,3)	-12,0 ***	(1,3)	-13,4 ***	(1,2)
„White collar“	16,8 ***	(0,9)	15,7 ***	(0,9)	20,8 ***	(0,8)
<i>Bücher im Haushalt (ausgelassen: keine)</i>						
1-10 Bücher	15,2 ***	(3,3)	14,0 ***	(3,1)	32,8 ***	(4,4)
11-50 Bücher	29,9 ***	(3,2)	29,7 ***	(2,9)	49,6 ***	(4,0)
51-100 Bücher	37,5 ***	(3,3)	37,6 ***	(3,0)	56,9 ***	(3,8)
101-250 Bücher	52,8 ***	(3,3)	51,4 ***	(3,0)	73,6 ***	(3,9)
251-500 Bücher	62,5 ***	(3,4)	62,1 ***	(3,1)	84,6 ***	(3,9)
Mehr als 500 Bücher	65,0 ***	(3,4)	63,6 ***	(3,1)	84,6 ***	(3,9)
<i>Ortslage der Schule (ausgelassen: weniger als 3.000 Einwohner)</i>						
Kleinstadt (3.000-15.000)	2,2	(2,5)	3,5	(2,4)	2,2	(2,6)
Stadt (15.000-100.000)	4,2	(2,7)	3,5	(2,6)	5,5 *	(2,8)
Großstadt (100.000-1.000.000)	6,2 **	(3,0)	5,1 *	(2,9)	8,5 ***	(3,1)
Zentrum von Großstadt mit > 1 Million Einw.	8,3 **	(3,6)	5,9 *	(3,4)	11,5 ***	(3,7)
Sonst. Lage in Großstadt mit > 1 Million E.	-2,0	(3,3)	1,2	(3,3)	4,6	(3,5)
Bruttoinlandsprodukt pro Kopf (1.000 \$)	-1,5 *	(0,8) ^a	-0,6	(0,8) ^a	0,2	(0,7) ^a

(weitergeführt auf der nächsten Seite)

Tabelle 1 (weitergeführt)

	Mathematik		Naturwissenschaften		Lesen	
	Koeffizient	Std.fehler	Koeffizient	Std.fehler	Koeffizient	Std.fehler
Häusliche Inputs						
<i>Elterliche Unterstützung (ausgelassen: fehlt ein wenig)</i>						
Fehlt sehr	-20,8 ***	(1,9)	-19,4 ***	(1,9)	-23,7 ***	(2,1)
Fehlt überhaupt nicht	15,0 ***	(2,6)	8,7 ***	(2,3)	10,5 ***	(2,5)
<i>Hausarbeiten</i>						
>1 und <3 Stunden pro Woche	8,6 ***	(0,8)	6,9 ***	(0,9)	8,4 ***	(0,7)
>3 Stunden pro Woche	11,1 ***	(1,1)	8,3 ***	(1,2)	4,4 ***	(1,1)
<i>Computer im Haushalt (ausgelassen: keiner)</i>						
Einer	-3,9 ***	(1,2)	-2,5 **	(1,1)	-4,2 ***	(1,0)
Mehr als einer	-9,8 ***	(1,2)	-10,7 ***	(1,2)	-10,4 ***	(1,1)
Schulressourcen und Lehrer						
Bildungsausgaben pro Schüler (1.000 \$)	10,7 ***	(3,6) ^a	7,7 **	(3,2) ^a	1,5	(4,6) ^a
Klassengröße (M/N/L) (instr. durch SLV)	0,8 *	(0,4)	1,6 ***	(0,4)	0,5	(0,4)
<i>Unterrichtsmaterialien (ausgelassen: fehlen ein wenig)</i>						
Fehlen überhaupt nicht	8,2 ***	(1,4)	8,1 ***	(1,3)	8,8 ***	(1,4)
Fehlen sehr	-13,9 ***	(3,5)	-6,1 *	(3,4)	-6,1	(3,8)
Unterrichtszeit (1.000 Minuten pro Jahr) (M/N/L)	0,8 ***	(0,2)	1,1 ***	(0,2)	-0,4 **	(0,2)
<i>Bildungsstand der Lehrer (Anteil in Schule)</i>						
„Master“ in Erziehungswissenschaften	4,3 *	(2,2)	9,4 ***	(2,0)	6,2 ***	(2,1)
Lehrerzertifikat	15,1 ***	(3,0)	15,2 ***	(3,0)	11,7 ***	(3,2)
„Masters“ im jeweiligen Fach (M/N/L)	10,9 ***	(2,6)	8,3 ***	(2,4)	16,4 ***	(2,8)
Institutionen						
<i>Prüfungen</i>						
Zentrale Abschlussprüfungen	19,1 *	(9,6) ^a	15,0	(9,1) ^a	6,9	(8,6) ^a
Standardisierte Tests	1,1	(1,5)	1,3	(1,4)	2,1	(1,5)
<i>Schulautonomie</i>						
Bestimmung der Lehrinhalte	1,9	(1,7)	-0,5	(1,6)	-1,1	(1,7)
Auswahl der Lehrbücher	20,0 ***	(4,6)	30,8 ***	(4,3)	31,3 ***	(5,2)
Formulierung des Schulbudgets	-7,5 ***	(1,8)	-6,4 ***	(1,7)	-3,1	(1,9)
Budgetverteilung innerhalb der Schule	12,6 ***	(3,2)	14,4 ***	(3,0)	10,1 ***	(3,0)
Einstellung von Lehrern	11,9 ***	(2,3)	2,7	(2,2)	5,3 **	(2,3)
Entlassung von Lehrern	-1,8	(2,1)	-2,7	(2,0)	-4,9 **	(2,2)
Festlegung der Anfangsgehälter der Lehrer	-3,0	(2,5)	-2,6	(2,2)	-4,9 **	(2,3)
Gehaltserhöhungen der Lehrer	-2,4	(2,4)	-3,5	(2,1)	-2,1	(2,3)
<i>Öffentliche vs. private Trägerschaft und Finanzierung</i>						
Öffentliche Trägerschaft	-19,5 ***	(2,4)	-15,8 ***	(2,2)	-15,4 ***	(2,2)
Öffentliche Finanzierung (Anteil)	13,5 ***	(3,8)	3,6	(3,4)	0,3	(3,5)
Schüler (Observationseinheiten)	96.855		96.758		174.227	
Schulen (Primäre Erhebungseinheiten)	6.611		6.613		6.626	
Länder (Schichten)	31		31		31	
R ² (ohne Dummies für imputierte Werte)	0,301		0,249		0,306	

Hinweise: Abhängige Variable: Punkte im internationalen PISA-Test.– 2SLS-Regression im jeweiligen Fach (Klassengröße instrumentiert mit dem Schüler-Lehrer-Verhältnis der Schule). Die Regressionen sind gewichtet mit den Erhebungswahrscheinlichkeiten der Schüler. Zur Vermeidung von Verzerrungen aufgrund fehlender Antworten in den Fragebögen enthalten die Regressionen als zusätzliche Kontrollvariablen Dummies für imputierte Werte für jede im Modell enthaltene Variable sowie für Interaktionen dieser Dummies mit der jeweiligen Variable. – Cluster-robuste Standardfehler in Klammern (berücksichtigen korrelierte Fehlerterme innerhalb der Schulen).– ^a Cluster-robuste Standardfehler (und damit Signifikanzniveaus) berücksichtigen korrelierte Fehlerterme innerhalb der Länder (statt Schulen).

Statistisches Signifikanzniveau (basierend auf Cluster-robusten Standardfehlern): *** 1%. – ** 5%. – * 10%.

Quelle: Fuchs und Wößmann (2004a).

Tabelle 2: Verfügbarkeit von Computern im Haushalt und PISA-Leistungen

	I	II	III	IV	V
<i>Computer im Haushalt (ausgelassen: keiner)</i>					
Einer	22,7*** (1,6)	17,1*** (1,5)	1,8 (1,3)	-2,1* (1,2)	-3,9*** (1,2)
Mehr als einer	29,5*** (1,6)	21,7*** (1,5)	-2,1 (1,4)	-6,6*** (1,3)	-9,8*** (1,2)
Schülereigenschaften	–	inkl.	inkl.	inkl.	inkl.
Familiärer Hintergrund	–	–	inkl.	inkl.	inkl.
Häusliche und schulische Inputs	–	–	–	inkl.	inkl.
Institutionen	–	–	–	–	inkl.
R^2	0,03	0,14	0,26	0,31	0,33
Schüler	96.855	96.855	96.855	96.855	96.855

Hinweise: Abhängige Variable: Punkte im internationalen PISA-Mathematik-Test. – 2SLS-Regressionen (Klassengröße instrumentiert mit dem Schüler-Lehrer-Verhältnis der Schule). Die Regressionen sind gewichtet mit den Erhebungswahrscheinlichkeiten der Schüler. Zur Vermeidung von Verzerrungen aufgrund fehlender Antworten in den Fragebögen enthalten die Regressionen als zusätzliche Kontrollvariablen Dummies für imputierte Werte für jede im Modell enthaltene Variable sowie für Interaktionen dieser Dummies mit der jeweiligen Variable. – Cluster-robuste Standardfehler in Klammern (berücksichtigen korrelierte Fehlerterme innerhalb der Schulen).

Statistisches Signifikanzniveau (basierend auf Cluster-robusten Standardfehlern): *** 1%. – ** 5%. – * 10%.

Quelle: Fuchs und Wößmann (2004b).

Ifo Working Papers

- No. 15 Flaig, G. and H. Rottmann, Labour Market Institutions and Employment Thresholds. An International Comparison, August 2005.
- No. 14 Hülsewig, O., E. Mayer and T. Wollmershäuser, Bank Loan Supply and Monetary Transmission in Germany: An Assessment Based on Matching Impulse Responses, August 2005.
- No. 13 Abberger, K., The Use of Qualitative Business Tendency Surveys for Forecasting Business Investing in Germany, June 2005.
- No. 12 Thum, M. Korruption und Schattenwirtschaft, Juni 2005.
- No. 11 Abberger, K., Qualitative Business Surveys and the Assessment of Employment – A Case Study for Germany, June 2005.
- No. 10 Berlemann, M. and F. Nelson, Forecasting Inflation via Experimental Stock Markets: Some Results from Pilot Markets, June 2005.
- No. 9 Henzel, S. and T. Wollmershäuser, An Alternative to the Carlson-Parkin Method for the Quantification of Qualitative Inflation Expectations: Evidence from the Ifo World Economic Survey, June 2005.
- No. 8 Fuchs, Th. and L. Wößmann, Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and Use of Computers at Home and at School, May 2005.
- No. 7 Werding, M., Survivor Benefits and the Gender Tax-Gap in Public Pension Schemes – Work Incentives and Options for Reform, May 2005.
- No. 6 Holzner, Chr., Search Frictions, Credit Constraints and Firm Financed General Training, May 2005.
- No. 5 Sülzle, K., Duopolistic Competition between Independent and Collaborative Business-to-Business Marketplaces, March 2005.

- No. 4 Becker, Sascha O., K. Ekholm, R. Jäckle and M.-A. Muendler, Location Choice and Employment Decisions: A Comparison of German and Swedish Multinationals, March 2005.
- No. 3 Bandholz, H., New Composite Leading Indicators for Hungary and Poland, March 2005.
- No. 2 Eggert, W. and M. Kolmar, Contests with Size Effects, January 2005.
- No. 1 Hanushek, E. and L. Wößmann, Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries, January 2005.