



Ursachenkomplexe der PISA-Ergebnisse: Untersuchungen auf Basis internationaler Mikrodaten¹

Ludger Wößmann

ifo Institut für Wirtschaftsforschung, Universität München

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.

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 zu Recht, lässt sich aus bildungsökonomischer Sicht hinzufügen. Mehrere Studien belegen, dass ein empirisch gesicherter positiver Zusammenhang zwischen den in internationalen Vergleichstests gemessenen Bildungsleistungen und der langfristigen wirtschaftlichen Entwicklung von Volkswirtschaften besteht.² Deshalb interessiert sich auch die bildungsökonomische Forschung in den letzten Jahren zunehmend für die Ursachen des Abschneidens in den internationalen Schülerleistungstests. Dabei werden zumeist drei Ursachen-

komplexe betrachtet: der familiäre Hintergrund der Schüler, die Ressourcenausstattung der Schulen und die institutionellen Rahmenbedingungen des Schulsystems.

Die vorliegende Studie berichtet über 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 der Schulen verbunden, welche detaillierte Informationen zu den drei Ursachenkomplexen liefern. Ein multivariates Herangehen ermöglicht, 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 würden alle Schulen den gleichen familiären Hintergrund aufweisen. Damit wird eine Konfundierung der jeweiligen Effekte mit anderen beobachteten Hintergrundeffekten ausgeschlossen. Die Berücksichtigung der internationalen Variation in den Schülerleistungen ermöglicht, 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 in und einen ersten Überblick über die zentralen Befunde gehen. Details zu methodischen Fragen, weiteren Ergebnissen, theoretischen Grundlagen und Interpretationen sowie ausführlichere Literaturverweise finden sich in Fuchs und Wößmann (2004a, 2004b) und Wößmann (2005a).³

Aus bildungspolitischer Sicht ist ein zentraler Befund der vorgestellten Arbeiten, dass die internationalen Unterschiede der 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.⁴ Entsprechend dieser theoretischen Basis belegen die hier vorgestellten empirischen Ergebnisse etwa, dass individuelle Schülerleistungen in Ländern besser sind, die zentrale Abschlussprüfungen durchführen, in Schulen, die Autonomie im Personalmanagement und in Prozessentscheidungen haben, und in Schulen, die in privater Trägerschaft sind. Demgegenüber zeigt sich kein eindeutiger Zusammenhang der Schülerleistungen mit dem Anteil der privaten Finanzierung der Schule, und ein negativer Zusammenhang zeigt sich mit Schulautonomie in Entscheidungsbereichen, die Raum bieten 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 Institutionen des Schulsystems legen nahe, dass die Schulpolitik durch

die institutionelle Gestaltung des Schulsystems durchaus weite Spielräume hat, auf die Leistungen der Schülerinnen und Schüler Einfluss zu nehmen.

Die Ergebnisse bestätigen ähnliche Befunde internationaler Mikroanalysen auf Schülerebene anhand früherer internationaler Schülerleistungstests (vgl. etwa Wößmann, 2001b, 2003a, b). 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, in Druck; Hanushek & Kimko, 2000; Lee & Barro, 2001). Demgegenüber können die hier präsentierten Mikrostudien lokale Unterschiede in den Hintergründen der Schüler wesentlich besser berücksichtigen und darüber hinaus mögliche Heterogenitäten systemischer Einflussfaktoren innerhalb der Schulsysteme abbilden.

Abschnitt 2 diskutiert die verwendete Methodik sowie die Daten. Abschnitt 3 berichtet zentrale Ergebnisse der Untersuchungen in Bezug auf die drei untersuchten Ursachenkomplexe familiärer Hintergrund, schulische Ressourcenausstattung und institutionelle Rahmenbedingungen. Dabei legt ein Exkurs zum Zusammenhang von Computern und Schülerleistungen die Notwendigkeit multivariater Vorgehensweisen dar. Ein 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 nachfolgend noch näher besprochen werden.⁵

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.⁶ Dabei werden die Daten aller an der PISA-Studie teilnehmenden Schülerinnen und 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.⁷

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 und Kulturen einheitlich sind. Mit der restriktiven Spezifikation ‚erkauft‘ man sich aber den Vorteil, dass die unüberschaubare Informationsflut der PISA-Daten auf ein Maß reduziert wird, welches generelle Muster (‚patterns‘) möglicher Wirkungszusammenhänge widerspiegeln kann. Obwohl es 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. Diese leistet im ersten Schritt eine Analyse *innerhalb* der Länder und untersucht dann im zweiten Schritt die verbleibende Variation *zwischen* den Ländern – die substantiellen Ergebnisse der beiden Modellspezifikationen unterscheiden sich nicht. 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 die hierarchische Struktur der Daten nach einer mehrbenenanalytischen Vorgehensweise. Dem trägt die vorliegende Untersuchung dadurch Rechnung, dass im Rahmen von cluster-robusten linearen Regressionen (CRLR) eine nicht-triviale Struktur der Fehlerterme abgebildet wird. In der PISA-Untersuchung stellen Schulen die primäre Erhebungseinheit dar, und es werden zahlreiche Variablen auf der Ebene der Schulen erhoben, so dass die individuellen Observationen innerhalb jeder Schule nicht als unabhängig voneinander betrachtet werden können. Deshalb hat der

Fehlerterm ε nicht nur die übliche Komponente auf der Ebene des individuellen Schülers ν_i , sondern zusätzlich 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 zugelassen wird, welcher der tatsächlich observierten Kovarianzstruktur der Kleinstquadrat-Fehlerterme entspricht (für Details vgl. Moulton, 1986; Deaton, 1997, S. 74–78).⁸ 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.⁹

Zusätzlich zu dieser Mehrebenenstruktur der Daten ist zu berücksichtigen, dass PISA ein stratifiziertes Erhebungsdesign innerhalb jedes 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 Stratifizierungseinheit mit einem zu ihrem Anteil an der Gesamtpopulation proportionalen Beitrag in die Schätzung eingeht (vgl. DuMouchel & 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 die Einflüsse des familiären Hintergrunds nicht kontrolliert, wird weiter unten am Beispiel des Einflusses von Computern auf die Schülerleistungen 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 stattdessen 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. Oft 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 PISA-Daten, die lediglich in einem Querschnitt vorliegen. 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 das gravierendere Endogenitätsproblem darstellen könnte (siehe West & Wößmann, 2005).¹⁰ Bei den anderen Ressourcenvariablen ist eine ähnliche Vorgehensweise aufgrund der Datenlage nicht umsetzbar. Die PISA-Daten bieten keine offensichtliche Möglichkeit, die verbleibenden Endogenitätsverzerrungen zu bewältigen. Daher können diese Ergebnisse auch nicht notwendigerweise kausal interpretiert werden, sondern sind als deskriptive Befunde zu bewerten (vgl. Fuchs & Wößmann, 2004a, b für weitere Details).¹¹

Wie angedeutet, lässt sich die restriktive Annahme der in allen Schulsystemen einheitlichen Effekte in einer erweiterten Spezifikation abschwächen. In Abschnitt 3.5 werden Ergebnisse einer Spezifikation berichtet, die 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:

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

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 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. Diese Basiskompetenzen bilden nur *ein* ‚Outputbündel‘ des Bildungsprozesses ab, und es sind sich wohl alle Beobachter einig, dass es noch weitere wichtige Bildungsziele innerhalb und außerhalb der Lehrpläne gibt. Allerdings wird auch kein Beobachter diesen Basiskompetenzen jegliche Bedeutung absprechen, und wie in der Einleitung berichtet, haben sie sich als wirtschaftlich höchst relevant erwiesen. Gleichzeitig wäre es 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.¹² 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-Publikationen (OECD, 2000, 2001, 2002; Adams & 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 zeigt Tabelle 1 (siehe Anhang). Die wichtigsten Ergebnisse werden im Folgenden nur beispielhaft innerhalb der drei Ursachenkomplexe familiärer Hintergrund, schulische Ressourcenausstattung und institutionelle Rahmenbedingungen angesprochen.

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.¹³ 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 & Minoiu, 2005), den Familienstatus,¹⁴ 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 Bildungshintergrunds 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.

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 (siehe Anhang)

beispielhaft für den Fall der Mathematikleistungen berichtet,¹⁵ so findet sich ein starker signifikanter positiver Zusammenhang.

Doch dieses bivariate Ergebnis ist trügerisch, wie die zusätzliche Berücksichtigung weiterer Einflussfaktoren zeigt. Der Zusammenhang wird schwächer, wenn man die Einflüsse der Variablen der Schülereigenschaften (vgl. Tab. 1) herausrechnet (Spalte II in Tab. 2). Nach Berücksichtigung der Einflüsse der Variablen des familiären Hintergrunds (vgl. Tab. 1) wird der Zusammenhang zwischen Computern und PISA-Leistungen sehr klein und statistisch insignifikant (Spalte III in Tab. 2). Und wenn 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, wie sehr einfache bivariate Betrachtungen irreführend sein können – auch wenn sie selbst in hochqualitativen Dokumenten wie den offiziellen PISA-Veröffentlichungen berichtet werden.¹⁶ Im vorliegenden Fall ist es offensichtlich, dass die Verfügbarkeit von Computern zu Hause stark von Bildungsniveau und Beruf der Eltern abhängt. Gut ausgebildete Eltern legen mehr Wert auf die Ausbildung ihrer Kinder. Daher verwundert nicht, 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 kein positiver, sondern sogar ein negativer Zusammenhang.

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

Neben weiteren Details berichten Fuchs und Wößmann (2004b) Ergebnisse zum Zusammenhang zwischen PISA-Leistungen und der Verfügbarkeit und Nutzung von Computern in der Schule: Bei multivariater Betrachtung ergibt sich 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 häufi-

geren 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. Tab. 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 pro Schüler) 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 & Wößmann, 2004a für Details).

In Bezug auf die Klassengröße findet sich kein negativer Zusammenhang zwischen Klassengröße und PISA-Leistungen (schlechtere Leistungen in größeren Klassen); in Mathematik und Naturwissenschaften zeigt sich vielmehr sogar ein signifikant positiver Zusammenhang (bessere Leistungen in größeren Klassen). Dies dürfte allerdings vor allem darauf zurückzuführen sein, dass vielerorts eine systematische Zuordnung schwächerer Schüler in Schulen mit kleineren Klassen stattfindet (vgl. Abschnitt 2.1 sowie West & Wößmann, 2005). Zusätzlich bleibt in der gewählten Spezifikation unberücksichtigt, dass es in unterschiedlichen Kulturen möglicherweise Unterschiede in den Effekten geben könnte.

Hingegen sind die PISA-Leistungen dort besser, wo geeignete Unterrichtsmaterialien (etwa Lehrbücher) vorhanden sind. In Mathematik und Naturwissenschaften ergibt sich auch ein positiver Zusammenhang mit der jährlichen Unterrichtszeit.¹⁷ 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 & 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.¹⁸ Während zentrale Abschlussprüfungssysteme zumeist nationale Merkmale sind,¹⁹ liegen im PISA-Datensatz auch Informationen auf Schulebene darüber vor, ob 15-jährige 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.²⁰

In Bezug auf die Autonomie von Schulen ergeben sich in verschiedenen Entscheidungsbereichen sehr unterschiedliche Ergebnisse.²¹ 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.²² 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.

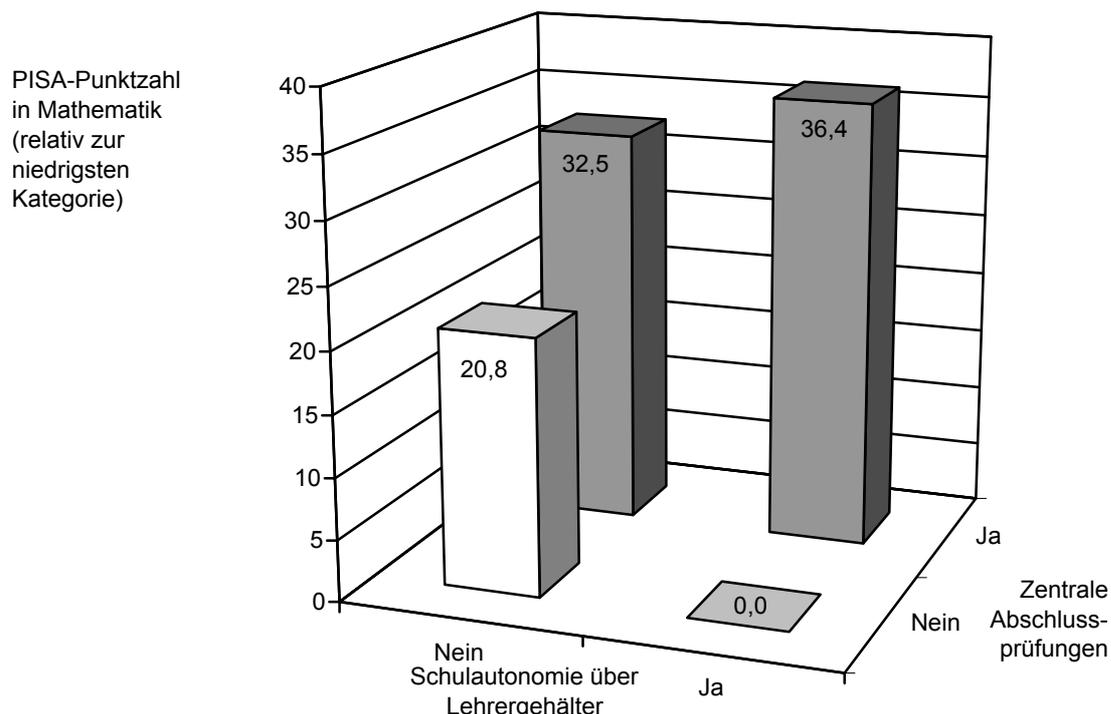
Schließlich liefert der PISA-Datensatz Informationen darüber, ob die jeweiligen Schulen in öffentlicher oder privater Trägerschaft operieren sowie wie hoch ihr öffentlicher und privater Finanzierungsanteil ist. Im Durchschnitt schneiden Schüler an Schulen in privater Trägerschaft in allen drei PISA-Kompetenzbereichen statistisch signifikant besser ab als Schüler an Schulen in öffentlicher Trägerschaft.²³ Aber sobald diese Effekte der Form der Trägerschaft berücksichtigt werden, besteht kein signifi-

kanter Zusammenhang zwischen dem Anteil der öffentlichen Finanzierung und den Schülerleistungen in Naturwissenschaften und Lesen. In Mathematik schneiden Schüler in Schulen mit einem höheren Anteil öffentlicher Finanzierung sogar besser ab.

3.5 Die Heterogenität des Effekts von externen Abschlussprüfungen

Das bisher betrachtete Modell erlaubte keine Heterogenität der Effekte zwischen verschiedenen Schulsystemen. Diese Restriktion wird nun ein Stück weit gelockert, 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). Beispielhaft wird auf besonders interessante Interaktionen zwischen zentralen Prüfungen und anderen institutionellen Gegebenheiten des Schulsystems, nämlich Schulautonomie und regelmäßige standardisierte Tests, eingegangen.

Abbildung 1: Zentrale Abschlussprüfungen und Gehaltsautonomie der Schulen



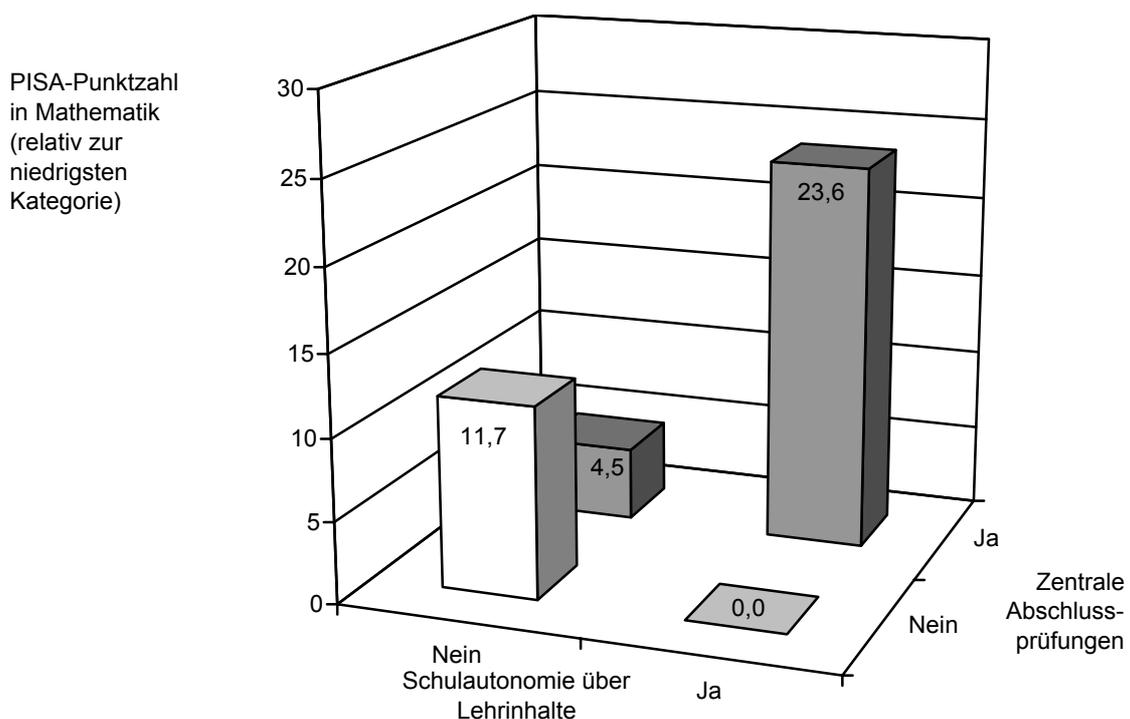
Quelle: Wößmann (2005a).

Aus theoretischer Sicht können solche Effektheterogenitäten erwartet werden, weil externe Prüfungen die im Bildungssystem bestehenden Informationsasymmetrien mildern können und mit dieser Transparenz ein gewisses Maß an Rechenschaft und Ver-

antwortlichkeit („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 stellen die Abbildungen 1–4 beispielhaft graphisch dar (siehe 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 Schulen, die Gehaltsautonomie haben. In Systemen ohne zentrale Prüfungen sind die PISA-Leistungen ebenfalls geringer in Schulen, die Eigenverantwortung über die Lehrinhalte haben (Abb. 2).

Abbildung 2: Zentrale Abschlussprüfungen und Schulautonomie über Lehrinhalte



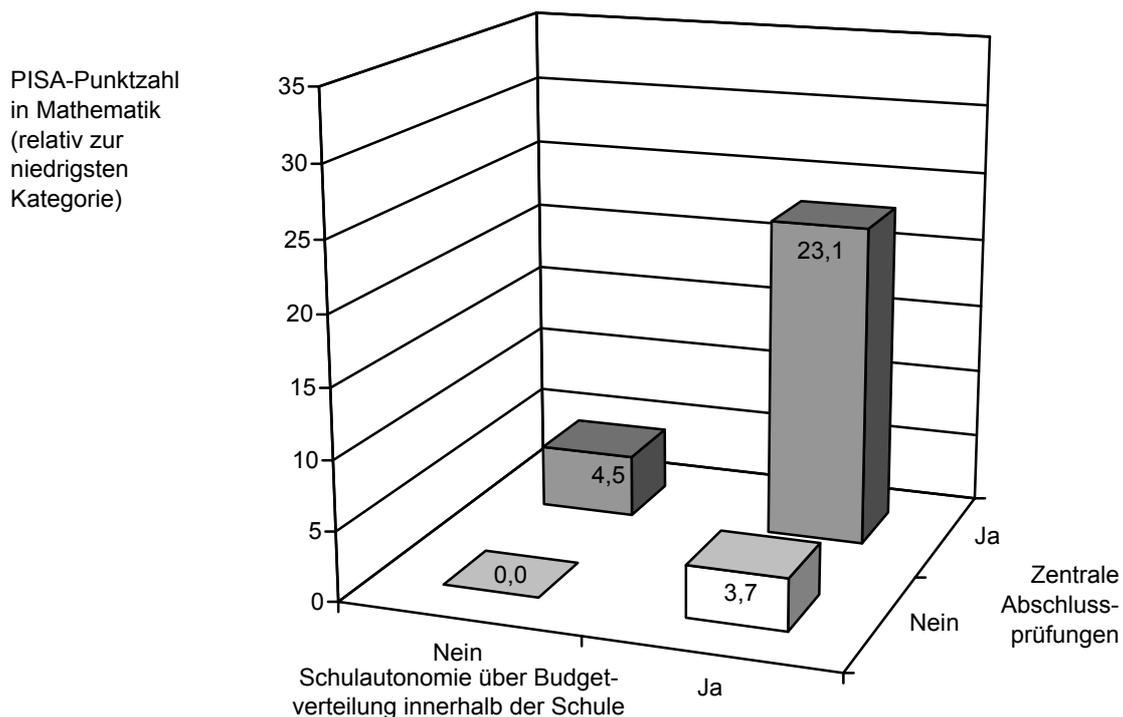
Quelle: Wößmann (2005a).

Dieser Zusammenhang dreht sich 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 zentrale 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 zentrale Abschlussprüfungen Verhaltensänderungen der schulischen Entscheidungsträger widerspiegeln.

Abbildung 3: Zentrale Abschlussprüfungen und Budgetverteilung innerhalb der Schulen

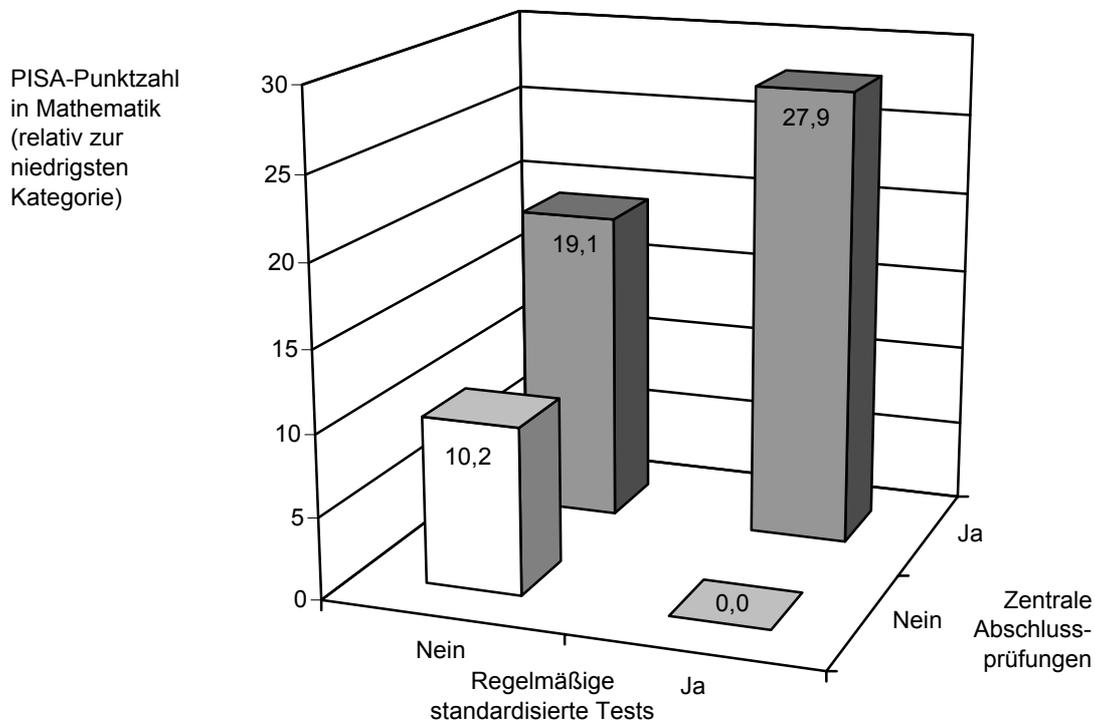


Quelle: Eigene Berechnungen anhand der PISA-Daten.

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 kön-

nen. 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.

Abbildung 4: Zentrale Abschlussprüfungen und regelmäßige standardisierte Tests



Quelle: Wößmann (2005a).

In Abbildung 4 zeigt sich schließlich ein ähnliches Bild für den Fall des regelmäßigen Einsatzes standardisierter Tests bei den 15-jährigen Schülern. Während in Systemen ohne zentrale Prüfungen ein negativer Zusammenhang zwischen den getesteten Schülerleistungen und standardisierten Tests besteht, ist dieser in Systemen mit zentralen Prüfungen signifikant positiv. Dies könnte darauf hindeuten, 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 erklärt. Fuchs und Wößmann (2004a) zeigen eine Methode, die den Erklärungsgehalt des Modells zwischen den Teilnehmerstaaten (im Unterschied zu zwischen den individuellen Schülern) abbildet. Dabei ergibt sich, dass das vorgestellte Modell 88 % der

zwischenstaatlichen Leistungsstreuung in Mathematik, 86 % in Naturwissenschaften und 85 % 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 lässt 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 erwartungsinkonforme positive Korrelation zwischen Leistungen und Klassengrößen bedingt ist (vgl. Fuchs & Wößmann, 2004a für weitere Details).

4. Schlussbemerkungen

Der vorliegende Beitrag berichtet 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-Teilnehmerstaaten lässt 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, Heijke und

Wößmann (2005), Wößmann (2005b, c), Wößmann und Fuchs (2005) und Wößmann und West (in Druck) von Interesse.

Auch wurden in den vorliegenden Untersuchungen aufgrund der gegebenen Datenbasis sicher 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 Lehrkräfte untersuchen (vgl. Wößmann, 2004a für einen Überblick).

Schließlich liegt der Fokus der vorgestellten Untersuchungen insbesondere auf bildungspolitischen Maßnahmen und dabei auf Fragen der effizienten Bildungsbereitstellung. 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 (in Druck) 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.

Anmerkungen

1. Der Autor dankt den Teilnehmerinnen und 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‘ sowie einem anonymen Gutachter für zahlreiche hilfreiche Kommentare und Diskussionen.
2. Vgl. Hanushek und Kimko (2000), Barro (2001), Wößmann (2002, 2003d) und Gundlach und Wößmann (2004).
3. Weitere methodische Details sowie Ergebnisse auf Basis weiterer internationaler Schülerleistungstests (TIMSS, TIMSS-Repeat und IGLU) finden sich in Wößmann (2001a, b, 2003a, b, c, e) und Wößmann und Fuchs (2005).
4. Vgl. Bishop und Wößmann (2004), Wößmann (2004a) und die Literaturverweise in Fuchs und Wößmann (2004a) für Details zur theoretischen Fundierung der hier analysierten Institutionen.
5. Zur Behandlung von in den Antwortfragebögen fehlenden Werten wurde eine auf geordneten Probit-Schätzungen basierende Imputationsmethode angewandt, die die weiteren im Text beschriebenen mikroökonomischen Feinheiten berücksichtigt. 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 Variablen, was zur Vermeidung von Verzerrungen aufgrund fehlender Werte beitragen soll; siehe Fuchs und Wößmann (2004a) für Details.
6. Für eine Einführung in die Grundmethodik der Regressionsanalyse vgl. jedes einschlägige Einführungslehrbuch in die Ökonometrie, z.B. Stock und Watson (2003).
7. Siehe etwa Hanushek (2002) für Details zur Spezifikation von Bildungsproduktionsfunktionen.

8. Bei Variablen, die nur auf der Ebene der Teilnehmerländer variieren, wird sogar jegliche vorhandene Kovarianz innerhalb der Länder in den Fehlertermen berücksichtigt.
9. In praktischer Hinsicht kommt die in der Ökonometrie häufiger verwendete cluster-robuste lineare Regression (CRLR) zumeist zu qualitativ identischen Ergebnissen als die verwandte, in der Erziehungswissenschaft häufiger verwendete Methode der Hierarchical Linear Models (HLM; vgl. etwa Bryk & 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.
10. Zusätzliche Verzerrungen der anhand der PISA-Daten geschätzten Ressourcenkoeffizienten können aufgrund von Aggregationsverzerrungen (Fertig & Wright, 2005) und aufgrund von über die Zeit variierenden Ausstattungsniveaus (vgl. Gundlach & Wößmann, 2000; Gundlach, Wößmann & Gmelin, 2001) auftreten.
11. Wößmann (2005c) und Wößmann und West (2005) untersuchen kausale Klassengrößeneffekte mit spezifischen Schätzmethode anhand der TIMSS-Daten.
12. Die Daten für Liechtenstein bleiben in den berichteten Untersuchungen aufgrund mangelnder international vergleichbarer Rahmendaten unberücksichtigt.
13. 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.
14. Für eine Analyse von Effekten der Geschwisterrivalität in PISA vgl. Wolter und Vellacott (2003).
15. Das Ergebnismuster in Naturwissenschaften und Lesen entspricht ganz dem in Mathematik.
16. Vgl. OECD (2001, S. 118), die – trotz Vorsichtsäußerungen bezüglich möglicher Interpretationsprobleme bivariater Zusammenhänge – folgenden bivariaten Befund berichtet: „Students with higher values on the index of interest in computers tend to perform better on the combined reading literacy scale.“
17. 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.
18. 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 (in Druck) folgend nur ein einfacher Durchschnitt des Vorliegens zentraler Prüfungen in Mathematik und Naturwissenschaften verwendet wurde.
19. Vgl. aber für Deutschland Jürges, Richter und Schneider (2004) und Jürges, Schneider und Büchel (2005), die ähnliche Befunde positiver Effekte von Zentralprüfungen zwischen den deutschen Bundesländern anhand der deutschen TIMSS- und PISA-E-Daten finden.
20. Ergebnisse von Quantilregressionen legen darüber hinaus nahe, dass der Effekt zentraler Abschlussprüfungen für gute Schüler etwas größer ist als für schwache Schüler (Wößmann, 2005a).
21. Vgl. Bishop und Wößmann (2004) für eine theoretische Erklärung dieser variierenden Ergebnisse zur Schulautonomie.

22. 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).
23. 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.

Literatur

- Adams, R. & Wu, M. (Eds.). (2002). *PISA 2000 technical report*. Paris: OECD.
- Akerhielm, K. (1995). Does class size matter? *Economics of Education Review*, 14 (3), 229–241.
- Ammermüller, A., Heijke, H. & Wößmann, L. (2005). Schooling quality in Eastern Europe: Educational production during transition. *Economics of Education Review*, 24 (5), 579–599.
- Barro, R. J. (2001). Human capital and growth. *American Economic Review, Papers and Proceedings*, 91 (2), 12–17.
- Bishop, J. H. (1997). The effect of national standards and curriculum-based exams on achievement. *American Economic Review*, 87 (2), 260–264.
- Bishop, J. H. (in Druck). Drinking from the fountain of knowledge: Student incentive to study and learn. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education*. Amsterdam: North-Holland.
- Bishop, J. H., & Wößmann, L. (2004). Institutional effects in a simple model of educational production. *Education Economics*, 12 (1), 17–38.
- Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and data analysis methods* (Advanced Quantitative Techniques in the Social Sciences, 1). Newbury Park: Sage.
- Cohen, J. & Baldi, S. (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, A. (1997). *The analysis of household surveys: A microeconomic approach to development policy*. Baltimore: The Johns Hopkins University Press.
- Dronkers, J. & Robert, P. (2003). *The effectiveness of public and private schools from a comparative perspective* (EUI Working Paper, SPS 2003-13). Florence: European University Institute.
- DuMouchel, W. H. & Duncan, G. J. (1983). Using sample survey weights in multiple regression analyses of stratified samples. *Journal of the American Statistical Association*, 78 (383), 535–543.
- Entorf, H. & Minoiu, N. (2005). What a difference immigration policy makes: A comparison of PISA scores in Europe and traditional countries of immigration. *German Economic Review*, 6 (3), 355–376.
- Fertig, M. & Wright, R. E. (2005). *School quality, educational attainment and aggregation bias* (Working-Paper-Version IZA Discussion Paper, 994). Verfügbar unter: <http://www.iza.org>. [Erscheint in *Economics Letters*].
- Fuchs, T. & Wößmann, L. (2004a). *What accounts for international differences in student performance? A re-examination using PISA data* (CESifo Working Paper, 1235). Munich: CESifo. Verfügbar unter: <http://www.cesifo.de>.

- Fuchs, T. & Wößmann, L. (2004b). Computers and student learning: Bivariate and multivariate evidence on the availability and use of computers at home and at school (CESifo Working Paper, 1321). Verfügbar unter: <http://www.cesifo.de>. [Erscheint in *Brussels Economic Review*, 47 (3)].
- Gundlach, E. & Wößmann, L. (2000). Die sinkende Produktivität der Schulen in OECD-Ländern. In M. Weiß & H. Weishaupt (Hrsg.), *Bildungsökonomie und Neue Steuerung* (Beiträge zur Bildungsplanung und Bildungsökonomie, Bd. 9) (S. 31–52). Frankfurt a.M.: Lang.
- Gundlach, E. & Wößmann, L. (2004). Bildungsressourcen, Bildungsinstitutionen und Bildungsqualität: Makroökonomische Relevanz und mikroökonomische Evidenz. In U. Backes-Gellner & P. Moog (Hrsg.), *Ökonomie der Evaluation von Schulen und Hochschulen* (Schriften des Vereins für Socialpolitik, N.F. Bd. 302) (S. 15–52). Berlin: Duncker & Humblot.
- Gundlach, E., Wößmann, L. & Gmelin, J. (2001). The decline of schooling productivity in OECD countries. *Economic Journal*, 111 (471), C135–C147.
- Hanushek, E. A. (2002). Publicly provided education. In A. J. Auerbach & M. Feldstein (Eds.), *Handbook of public economics, Volume 4* (pp. 2045–2141). Amsterdam: North Holland.
- Hanushek, E. A. & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review*, 90 (5), 1184–1208.
- Hanushek, E. A. & Wößmann, L. (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: <http://www.nber.org>. [Erscheint in *Economic Journal*].
- Jürges, H., Richter, W. F. & Schneider, K. (2004). Teacher quality and incentives: Theoretical and empirical effects of standards on teacher quality (CESifo Working Paper, 1296). Munich: CESifo. Verfügbar unter: <http://www.cesifo.de>.
- Jürges, H., Schneider, K. & Büchel, F. (2005). The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany. *Journal of the European Economic Association*, 3 (5), 1134–1155.
- Lee, J.-W. & Barro, R. J. (2001). Schooling quality in a cross-section of countries. *Economica*, 68 (272), 465–488.
- Moulton, B. R. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics*, 32 (3), 385–397.
- OECD (Organisation for Economic Co-operation and Development) (2000). *Measuring student knowledge and skills: The PISA 2000 assessment of reading, mathematical and scientific literacy*. Paris: OECD.
- OECD (Organisation for Economic Co-operation and Development) (2001). *Knowledge and skills for life: First results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris: OECD.
- OECD (Organisation for Economic Co-operation and Development) (2002). *Manual for the PISA 2000 database*. Paris: OECD.
- Peterson, P. & Wößmann, L. (Eds.). (in Druck). *Schools and the equal opportunity problem*. Cambridge, MA: MIT Press.
- Schütz, G. & Wößmann, L. (2005). *Chancengleichheit im Schulsystem: Internationale deskriptive Evidenz und mögliche Bestimmungsfaktoren* (Ifo Working Paper, 17). München: ifo Institut für Wirtschaftsforschung an der Universität München. Verfügbar unter: <http://www.cesifo.de>.
- Stock, J. H. & Watson, M. W. (2003). *Introduction to econometrics*. Boston: Pearson Education / Addison-Wesley.
- West, M. R. & Wößmann, L. (2005). Which school systems sort weaker students into smaller classes? International evidence (CESifo Working Paper, 1054) Verfügbar unter: <http://www.cesifo.de>. [Erscheint in *European Journal of Political Economy*].

- White, H. (1984). *Asymptotic theory for econometricians*. Orlando: Academic Press.
- Wirth, J. & Klieme, E. (2003). Computernutzung. In Deutsches PISA-Konsortium (Hrsg.), *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 195–209). Opladen: Leske + Budrich.
- Wößmann, L. (2001a). Why students in some countries do better: International evidence on the importance of education policy. *Education Matters*, 1 (2), 67–74.
- Wößmann, L. (2001b). Schulsystem und Schülerleistung im internationalen Vergleich: Was Institutionen ausmachen. *Die Weltwirtschaft*, (3), 283–304.
- Wößmann, L. (2002). *Schooling and the quality of human capital*. Berlin: Springer.
- Wößmann, L. (2003a). Schooling resources, educational institutions and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics*, 65 (2), 117–170.
- Wößmann, L. (2003b). Central exit exams and student achievement: International evidence. In P. E. Peterson & M. R. West (Eds.), *No child left behind? The politics and practice of school accountability*. Washington, D.C.: Brookings Institution Press.
- Wößmann, L. (2003c). Zentrale Prüfungen als „Währung“ des Bildungssystems: Zur Komplementarität von Schulautonomie und Zentralprüfungen. *Vierteljahreshefte zur Wirtschaftsforschung*, 72 (2), 220–237.
- Wößmann, L. (2003d). Specifying human capital. *Journal of Economic Surveys*, 17 (3), 239–270.
- Wößmann, L. (2003e). Familiärer Hintergrund, Schulsystem und Schülerleistungen im internationalen Vergleich. *Aus Politik und Zeitgeschichte*, B 21–22, 33–38.
- Wößmann, L. (2004a). Institutional comparisons in educational production. *CESifo DICE Report – Journal for Institutional Comparisons*, 2 (4), 3–6.
- Wößmann, L. (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: <http://www.cesifo.de>.
- Wößmann, L. (2005a). The effect heterogeneity of central exams: Evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics*, 13 (2), 143–169.
- Wößmann, L. (2005b). Educational production in East Asia: The impact of family background and schooling policies on student performance. *German Economic Review*, 6 (3), 331–353.
- Wößmann, L. (2005c). Educational production in Europe. *Economic Policy*, 20 (43), 445–504.
- Wößmann, L., unter Mithilfe von T. 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: <http://www-wds.worldbank.org>.
- Wößmann, L. & West, M. R. (in Druck). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*.
- Wolter, S. C. & Vellacott, M. C. (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, J. M. (2001). Asymptotic properties of weighted *M*-Estimators for standard stratified samples. *Econometric Theory*, 17 (2), 451–470.

Anhang

Tabelle 1: Internationale ‚Bildungsproduktionsfunktionen‘

	Mathematik		Naturwissenschaften		Lesen	
	Koeffizient	Std.fehler	Koeffizient	Std.fehler	Koeffizient	Std.fehler
<i>Schülereigenschaften</i>						
<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)
<i>Familiärer Hintergrund</i>						
<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</i>						
<i>(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</i>						
<i>(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</i>						
<i>(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</i>						
<i>(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</i>						
<i>(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)

Tabelle 1: Internationale ‚Bildungsproduktionsfunktionen‘ Fortsetzung

	Mathematik		Naturwissenschaften		Lesen	
	Koeffizient	Std.fehler	Koeffizient	Std.fehler	Koeffizient	Std.fehler
<i>Familiärer Hintergrund</i>						
<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 E.	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
<i>Häusliche Inputs</i>						
<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)
<i>Schulressourcen und Lehrer</i>						
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)
<i>Institutionen</i>						
<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)

Tabelle 1: Internationale ‚Bildungsproduktionsfunktionen‘ Fortsetzung

	Mathematik		Naturwissenschaften		Lesen	
	Koeffizient	Std.fehler	Koeffizient	Std.fehler	Koeffizient	Std.fehler
<i>Institutionen</i>						
<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	

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 Variablen.

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).

*** 1 %; ** 5 %; * 10 % Statistisches Signifikanzniveau (basierend auf cluster-robusten Standardfehlern).
Quelle: Fuchs und Wößmann (2004a).

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

	I	II	III	IV	V
<i>Computer im Haushalt</i>					
<i>(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

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 Variablen.

Cluster-robuste Standardfehler in Klammern (berücksichtigen korrelierte Fehlerterme innerhalb der Schulen).

*** 1 %; ** 5 %; * 10 % Statistisches Signifikanzniveau (basierend auf cluster-robusten Standardfehlern).

Quelle: Fuchs und Wößmann (2004b).