

Working Papers

Chancengleichheit im Schulsystem: Internationale deskriptive Evidenz und mögliche Bestimmungsfaktoren

Gabriela Schütz
Ludger Wößmann

Ifo Working Paper No. 17

Oktober 2005

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

Chancengleichheit im Schulsystem: Internationale deskriptive Evidenz und mögliche Bestimmungsfaktoren*

Abstract

Mikroökonomische Schätzungen anhand internationaler Schülerleistungstests zeigen, dass das Ausmaß, in dem Schülerleistungen mit dem familiären Hintergrund variieren, international sehr unterschiedlich ist. So weisen etwa Frankreich und Kanada wesentlich ausgeglichene Bildungschancen für Kinder unterschiedlicher familiärer Herkunft auf als Deutschland und Großbritannien. Verschiedene empirische Identifikationsstrategien belegen, dass Schulsysteme mit späterer Mehrgliedrigkeit und einem umfassenderen Vorschulsystem systematisch ausgeglichene Bildungschancen bieten. Dagegen findet sich kein systematischer Zusammenhang der Chancengleichheit mit Ganztagschulsystemen, dem Ausgabenniveau oder der Durchschnittsleistung.

JEL Code: I28, D63.

Keywords: Chancengleichheit, Bildungsökonomik, Schülerleistungen, TIMSS, familiärer Hintergrund, soziale Selektivität, Mehrgliedrigkeit, Vorschulsystem.

Gabriela Schütz
ifo Institut für Wirtschaftsforschung
an der Universität München und CESifo
Poschingerstr. 5
81679 München
Tel.: +49(0)89/9224-1698
E-Mail: schuetz@ifo.de
Intenet: www.cesifo-group.de/link/25IFOHS

Ludger Wößmann
ifo Institut für Wirtschaftsforschung
an der Universität München und CESifo
Poschingerstr. 5
81679 München
Tel.: +49(0)89/9224-1699
E-Mail: woessmann@ifo.de
Intenet: www.cesifo-group.de/link/25IFOHS

* Wir möchten uns bei den Teilnehmern der Jahrestagung des Bildungsökonomischen Ausschusses des Vereins für Socialpolitik 2005 in Nürnberg für hilfreiche Kommentare und Diskussionen herzlich bedanken.

A. Einleitung

Während sich die empirische bildungsökonomische Forschung auf dem Gebiet der Bildungsproduktion bisher weitgehend auf Effizienzfragen konzentriert hat, stehen in der politischen Diskussion regelmäßig auch Distributionsfragen im Mittelpunkt. Internationale Vergleichsanalysen legen darüber hinaus nahe, dass es gerade Unterschiede in der Streuung der Bildungsleistungen sind, die – weit mehr als etwa Mindestlöhne oder gewerkschaftlicher Organisationsgrad – geeignet sind, die existierenden Unterschiede in der Einkommensungleichheit zwischen Ländern zu erklären (siehe Nickell 2004).

Der vorliegende Beitrag nutzt deshalb zunächst die Mikrodaten der TIMSS- und der TIMSS-Repeat-Studien, um multivariat zu schätzen, wie stark in den 54 teilnehmenden Ländern die Bildungsleistungen mit dem familiären Hintergrund variieren (Abschnitt B). Die Stärke des familiären Einflusses dient zugleich als ein international vergleichbarer Proxy für die Chancengleichheit in der Bildung. Dabei ergibt sich – ähnlich wie auch aus der Diskussion der PISA-Ergebnisse bekannt –, dass Länder wie Großbritannien und Deutschland eine relativ geringe Chancengleichheit für Kinder unterschiedlicher sozialer Herkunft aufweisen, während Länder wie Frankreich und Kanada deutlich ausgeglichene Leistungen erzielen.

Darauf aufbauend wird möglichen bildungspolitischen Ursachen dieser internationalen Unterschiede in der Chancengleichheit nachgegangen. Eine erste Identifikationsstrategie interagiert in dem internationalen Mikrodatensatz Merkmale der nationalen Bildungspolitik mit dem familiären Hintergrund auf Schülerebene (Abschnitt C.I). Die Ergebnisse zeigen, dass sowohl eine späte Mehrgliedrigkeit als auch ein umfassendes Vorschulsystem mit ausgeglicheneren Bildungschancen für Kinder unterschiedlicher familiärer Herkunft einhergehen. Demgegenüber scheinen das Ausgabeniveau, das durchschnittliche Leistungsniveau, die Länge des Schultages und das wirtschaftliche Entwicklungsniveau in keinem signifikanten systematischen Zusammenhang zur Chancengleichheit zu stehen.

Eine zweite Identifikationsstrategie befasst sich gezielt mit möglichen Effekten der frühen Mehrgliedrigkeit (Abschnitt C.II). Um Verzerrungen durch unbeobachtete Heterogenität zwischen den Ländern auszuschließen, verwendet diese Strategie einen Differenzen-in-Differenzen-Schätzer, der die Differenz in der Ungleichheit zwischen dem Grund- und dem Sekundarschulbereich zwischen Ländern mit und ohne frühe Mehrgliedrigkeit vergleicht. Die Ergebnisse aus acht Paaren verschiedener internationaler Schülerleistungstests deuten darauf hin, dass eine frühe Mehrgliedrigkeit in der Tat die Streuung der Bildungsleistungen signifikant erhöht. Obwohl weniger eindeutig, scheint eine frühe Mehrgliedrigkeit zudem das durch-

schnittliche Leistungsniveau eher zu senken als zu erhöhen, was einen Zielkonflikt zwischen Gleichheit und Effizienz in diesem Bereich ausschließen würde.

B. Internationaler Vergleich der Chancengleichheit im Schulsystem

Dieser Abschnitt widmet sich dem deskriptiven Ziel aufzuzeigen, inwiefern verschiedene Schulsysteme allen Kindern unabhängig von ihrem sozioökonomischen Hintergrund gleiche Bildungschancen bieten.

1. Familiärer Hintergrund und Schülerleistungen in Europa und den USA

Die in einem Schulsystem bestehende Chancenungleichheit zwischen Kindern unterschiedlicher familiärer Herkunft lässt sich mit Hilfe der Daten internationaler Schülerleistungsstudien als die Stärke des Einflusses des familiären Hintergrundes auf die Testleistung empirisch schätzen. Damit können die Schulsysteme als ihren Schülern im internationalen Vergleich relativ viel oder relativ wenig Chancengleichheit bietend beschrieben werden.

Wir beginnen mit einer aggregierten Analyse dieser Zusammenhänge für Europa und die USA anhand der Daten der Third International Mathematics and Science Study (TIMSS), die 1995 unter Federführung der International Association for the Evaluation of Educational Achievement (IEA) durchgeführt wurde. In TIMSS wurden national repräsentative Stichproben der beiden Jahrgangsstufen mit dem größten Anteil 13jähriger Schüler (zumeist die 7. und 8. Klasse) in Mathematik und Naturwissenschaften getestet. Im Rahmen von TIMSS wurden nicht nur die individuellen Leistungen getestet, sondern über einen getrennten Schülerfragebogen außerdem Informationen zum familiären Hintergrund der Schüler erfasst. Diese Kombination aus Leistungstest und Hintergrundinformation auf Schülerebene erlaubt es, eine Bildungsproduktionsfunktion der folgenden Form zu schätzen:

$$T_{isc} = \alpha + F_{isc} \beta_1 + (F_{isc} E_c) \beta_2 + C_c \delta + \varepsilon_{isc} , \quad (1)$$

wobei T_{isc} die Testpunktzahl des Schülers i in Schule s in Land c und F_{isc} einen Vektor von Variablen des familiären Hintergrundes bezeichnet. Die Regression enthält eine Konstante α und einen Dummy für europäische Länder E_c . Damit gibt in dieser Schätzgleichung der Kleinst-Quadrate-Schätzer von β_1 die Effekte des Familienhintergrundes in den USA wieder, und der Schätzer von β_2 den Unterschied dieser Effekte zwischen den USA und den europäi-

schen Ländern. Zusätzlich kontrolliert die Schätzgleichung für Niveauheterogenität zwischen den betrachteten Ländern, indem sie für einen Satz von Länderdummies C_c kontrolliert.¹

Bei der Schätzung des Fehlerterms ε_{isc} muss die zweistufige Stichprobenerhebung des Tests berücksichtigt werden (vgl. Moulton 1986 für dieses Problem der hierarchischen Datenstruktur). In TIMSS wurden innerhalb eines jeden Landes zuerst Schulen zufällig ausgewählt, und in einem zweiten Schritt Klassen innerhalb dieser Schulen (vgl. Martin und Kelly 1996). Die primäre Erhebungseinheit in TIMSS ist also die Schule. Daher schätzen wir die Gleichungen im gesamten Beitrag mit der Methode der cluster-robusten linearen Regression (CRLR), die dieses Stichprobendesign berücksichtigt, indem sie in der Varianz-Kovarianz-Matrix der Fehlerterme Interdependenzen innerhalb einer primären Erhebungseinheit zulässt und nur annimmt, dass die Beobachtungen zwischen den primären Erhebungseinheiten unabhängig sind (vgl. White 1984; Deaton 1997). Die Regressionsgleichung berücksichtigt zudem im Rahmen einer gewichteten Kleinst-Quadrate-Schätzung (WLS) die unterschiedlichen Erhebungswahrscheinlichkeiten verschiedener Schüler, so dass jede Teilpopulation mit einem zu ihrem Anteil an der Gesamtpopulation proportionalen Beitrag in die Schätzung eingeht (vgl. DuMouchel und Ducan 1983; Wooldridge 2001).²

Tabelle 1 berichtet die Ergebnisse einer Regression von Gleichung (1) für die 17 in TIMSS teilnehmenden westeuropäischen Schulsysteme³ und die USA. Dabei beinhaltet die erste Spalte die für die USA geltenden Koeffizienten β_1 , während die zweite Spalte die Effekte β_2 der Interaktion zwischen den einzelnen Kontrollvariablen und einem Europadummy und damit den Unterschied der Effekte zwischen den USA und den europäischen Ländern wiedergibt.

Die Ergebnisse der Regression belegen zahlreiche statistisch und quantitativ signifikante familiäre Einflüsse auf die Schülerleistung. Kinder, deren Eltern einen Universitätsabschluss besitzen, erzielen in den USA z.B. im Durchschnitt 52 Testpunkte mehr als Kinder, deren Eltern über keine Sekundarbildung verfügen.⁴ Dies ist sogar noch etwas mehr als der Vor-

¹ Zusätzlich kontrolliert die Schätzgleichung für Imputationsdummies und deren Interaktionen mit den Originalvariablen, da der in diesem Unterabschnitt benutzte Datensatz imputierte Werte für fehlende Antworten enthält (vgl. Wößmann 2004 für Details).

² Für weitere Details zu den verwendeten Methoden und Daten siehe Wößmann (2004).

³ Im Folgenden benutzen wir die Begriffe System und Land synonym, wobei zu berücksichtigen ist, dass England und Schottland sowie das flämische und französische Schulsystem in Belgien jeweils separat teilnehmen.

⁴ Bei der Variable „Bildung der Eltern“ wurde der jeweils höhere Bildungsstand der beiden Elternteile berücksichtigt.

sprung von 46 Punkten, den Kinder dadurch erzielen, dass sie zu der höheren der beiden getesteten Klassenstufen gehören. Da die Stichprobe derart standardisiert wurde, dass die mittlere Testpunktzahl 500 und die internationale Standardabweichung 100 beträgt, lassen sich die Koeffizientenschätzer direkt als Prozent einer internationalen Standardabweichung interpretieren.

Der Zusammenhang zwischen Schülerleistungen und den zu Hause verfügbaren Büchern ist sogar noch stärker. Die Anzahl der im Haushalt des Schülers verfügbaren Bücher lässt sich als ein allgemeiner Proxy für den sozioökonomischen Hintergrund der Familie interpretieren. Eine große Anzahl von Büchern kann als ein Indikator für eine familiäre Umgebung gelten, die schulischen Erfolg und Bildung hochschätzt und deshalb auch die akademischen Bemühungen ihrer Kinder fördert. Die Ergebnisse zeigen, dass ein Kind, dessen Familie mehr als 200 Bücher besitzt, im Durchschnitt 62 Punkte mehr erzielt als ein Kind, dessen Familie weniger als 11 Bücher besitzt – sowohl in den USA als auch in Europa.

Als zusätzliche Maße des familiären Hintergrundes werden in der Regression der Familienstatus (d.h., ob der Schüler mit beiden Elternteilen zusammen lebt), der Immigrationsstatus (d.h., ob der Schüler in dem Land geboren wurde), das Geschlecht und das Alter der Schüler sowie die geographische Lage der Schule berücksichtigt. Insgesamt ergeben sich in allen Ländern sehr starke familiäre Einflüsse: Kinder mit „gutem“ familiärem Hintergrund erzielen im Durchschnitt weit bessere Leistungen als ihre vergleichbaren Klassenkameraden mit „schlechterem“ Hintergrund. Weitere Studien anhand anderer Schülerleistungstests zeigen übrigens, dass dieser Befund unabhängig davon gilt, welche Fähigkeiten getestet werden (z.B. Leseverständnis, Mathematik oder Naturwissenschaften) und welches Maß für den familiären Hintergrund verwendet wird (z.B. Bildung oder Einkommen der Eltern, Anzahl der im Haushalt vorhandenen Bücher oder Indices für den sozioökonomischen Status der Familie).⁵

Der Vergleich zwischen erster und zweiter Spalte in Tabelle 1 zeigt außerdem, dass die Größenordnung der familiären Hintergrundeffekte in den USA und Europa erstaunlich ähnlich sind. Dies ist allerdings nicht als Beleg für ein international einheitliches Ausmaß von Chancengleichheit in den betreffenden Schulsystemen zu werten. Vielmehr verschleiert die aggregierte Betrachtung die zwischen den europäischen Ländern tatsächlich bestehenden

⁵ So finden beispielsweise Noël und de Broucker (2001) auf der Grundlage der Mikrodaten der International Adult Literacy Survey (IALS) einen starken Einfluss des Bildungsniveaus der Eltern auf die Anzahl der Schuljahre, welche die Teilnehmer absolviert haben, und auf die im Test erreichte Lesekompetenz. Ebenso weisen Baumert und Schümer (2001) auf der Grundlage der Daten der PISA-2000-Studie einen starken und statistisch signifikanten Zusammenhang zwischen sozialer Herkunft und Kompetenzerwerb nach.

großen Unterschiede (vgl. dazu im Detail Wößmann 2004). Im nächsten Abschnitt werden wir uns deshalb auf eine länderspezifische Betrachtungsweise konzentrieren. Um die Betrachtung fokussieren zu können, beschränken wir die Analyse außerdem auf einen einzigen international vergleichbaren Proxy für den sozioökonomischen Hintergrund der Familie.

II. Ein Maß der Chancenungleichheit in 54 Ländern

Im Folgenden erweitern wir die zuvor genutzte Stichprobe um zusätzliche Länder, indem wir die Daten der TIMSS-Studie von 1995 mit den Daten der TIMSS-Repeat-Studie zusammenfassen, die 1999 in zum Teil anderen Ländern durchgeführt wurde. Dadurch können wir die Analyse für insgesamt 54 Länder vornehmen (vgl. Schütz et al. 2005 für Details zu diesem Datensatz).

Als international vergleichbaren Proxy für den sozioökonomischen Familienhintergrund nutzen wir in der weiteren Analyse die Anzahl der im Haushalt des Schülers verfügbaren Bücher. Die Bücheranzahl ist ein in der soziologischen Literatur vorgeschlagener und vielfach verwendeter Indikator für den familiären Hintergrund. Aus mehreren Gründen halten wir ihn für präferabel im Vergleich zu alternativen Maßen, wie insbesondere dem elterlichen Bildungsstand. Erstens hat sich die Variable Anzahl der Bücher, wie schon in Tabelle 1 belegt, regelmäßig als stärkster einzelner Prädiktor für die Schülerleistung erwiesen. Dieser Befund ergibt sich nicht nur in TIMSS (Wößmann 2003; 2004) und TIMSS-Repeat (Wößmann 2002), sondern etwa auch in PISA (Fuchs und Wößmann 2004). Zweitens unterscheiden sich spezifische Bildungsniveaus oft stark zwischen Ländern, auch wenn sie mit Hilfe der International Standard Classification of Education (ISCED) berichtet werden. Daher impliziert dasselbe elterliche Bildungsniveau in verschiedenen Ländern manchmal unterschiedliche Wissensniveaus. Drittens bestehen bei der Variable Bücheranzahl weit geringere Probleme mit fehlenden Antworten in den Fragebögen als bei anderen möglichen Maßen des familiären Hintergrundes. So fehlen in TIMSS z.B. im Durchschnitt der westeuropäischen Länder weniger als 3 Prozent der Beobachtungen für die Variable Anzahl der Bücher, verglichen mit circa einem Drittel der Beobachtungen für die Variable Bildung der Eltern. Viertens besteht in TIMSS noch ein zusätzliches spezifisches Datenproblem, da das Maß der elterlichen Bildung zwischen der 1995er und der 1999er Ausgabe des Tests aufgrund von Anpassungen in der ISCED-Klassifikation leicht differiert, während die Frage zur Bücheranzahl in identischer Weise gestellt wurde. Aus diesen Gründen erscheint es sinnvoll, die Bücher-Variable in der vorliegenden Studie anderen Maßen des familiären Hintergrundes vorzuziehen, wie z.B. der

Bildung, dem Berufsstand oder dem Einkommen der Eltern oder auch Indices des sozioökonomischen Status der Familie, die aus den zuvor genannten Informationen gebildet werden.⁶

Um das Ausmaß der Chancenungleichheit innerhalb der Schulsysteme zu schätzen und international vergleichbar zu machen, regressieren wir für jedes Land c die Testpunktzahl der einzelnen Schüler auf unsere Proxyvariable für den familiären Hintergrund (Bücher im Haushalt B), eine Anzahl von Kontrollvariablen F_2 und einen Dummy für die jeweilige Studie S (TIMSS vs. TIMSS-Repeat):

$$T_{isc} = \alpha_c + \beta_c B_{isc} + F_{2,isc} \gamma_c + \delta_c S_{isc} + \varepsilon_{isc} . \quad (2)$$

Die Reihe der Kontrollvariablen F_2 umfasst wiederum das Alter und Geschlecht des Schülers und einen Dummy für den Familienstatus. In Bezug auf den Immigrationsstatus geht sie sogar über Gleichung (1) hinaus, indem sie nicht nur einen Dummy für den Immigrationsstatus des Schülers beinhaltet, sondern auch noch jeweils einen für den Immigrationsstatus der Mutter und des Vaters. Schließlich umfasst F_2 auch noch Interaktionsterme zwischen diesen drei Immigrationsdummies und der Bücher-Variablen.

Der Koeffizient β_c an der Variable B_{isc} ist das von uns gesuchte Maß für die Chancenungleichheit in dem betreffenden Land c , weil er die Stärke des Einflusses des familiären Hintergrundes (EFH) auf die Schülerleistung schätzt. Die Anzahl der im Haushalt des Schülers verfügbaren Bücher B_{isc} ist dabei eine kategoriale Variable, welche Werte zwischen 1 und 5 annimmt in Übereinstimmung mit den fünf Antwortkategorien auf die Frage „Wie viele Bücher gibt es bei Dir zu Hause (ausgenommen Zeitschriften, Zeitungen und Schulbücher)?“: „keine oder sehr wenige (0-10 Bücher)“; „genug um ein Regalbrett zu füllen (11-25 Bücher)“; „genug um ein Bücherregal zu füllen (26-100 Bücher)“; „genug um zwei Bücherregale zu füllen (101-200 Bücher)“; „genug um drei oder mehr Bücherregale zu füllen (mehr als 200 Bücher)“. Diese funktionale Form, die annimmt, dass die Leistungsunterschiede von Schülern zwischen je zwei aufeinander folgenden Kategorien gleich ist, hat sich in Spezifikationstests als den Daten entsprechend erwiesen.

Im Gegensatz zu den offiziellen Veröffentlichungen im Zusammenhang mit TIMSS (vgl. Beaton et al. 1996; Mullis et al. 2000) bestehen unsere Analysen nicht in einfachen bivariaten Vergleichen der durchschnittlichen Leistung von Schülern innerhalb verschiedener Katego-

⁶ Informationen über den Berufsstand oder das Einkommen der Eltern liegen in den TIMSS-Tests nicht vor. Als Beispiel für einen Index basiert der vielfach betrachtete International Socio-Economic Index (ISEI) auf einer Klassifikation des Berufsstatus, der auf einer älteren Studie von 16 Ländern beruht und dessen zugrunde liegende Annahmen daher fragwürdig und vielfach auch unklar erscheinen.

rien einer familiären Hintergrundvariablen. Stattdessen kontrolliert unsere Regressionsgleichung für die Effekte weiterer potentiell wichtiger Einflussvariablen (wie Alter, Geschlecht, Familienzusammensetzung und Immigrationshintergrund) und hält dadurch diese Faktoren bei der Analyse des Einflusses der Büchervariable konstant.

Besonders die Tatsache, dass die betrachteten Länder unterschiedliche Immigrationspopulationen aufweisen, könnte die Ergebnisse bivariater Schätzungen des Einflusses des Familienhintergrundes verzerren. Gegeben, dass Immigrationsstatus und familiärer Hintergrund in vielen Ländern stark korreliert sind, sind die internationalen Unterschiede in den Regressionsergebnissen für den Einfluss des familiären Hintergrundes verzerrt, sofern der Immigrationsstatus bei der Schätzung nicht berücksichtigt wird. Die drei Dummies für den Immigrationsstatus des Schülers, seiner Mutter und seines Vaters dienen der Verhinderung einer solchen Verzerrung. Darüber hinaus könnte der Einfluss des familiären Hintergrundes heterogen sein zwischen im Land geborenen und eingewanderten Familien, was ebenfalls zu einer Verzerrung der Schätzergebnisse führen würde. Um eine solche Verzerrung zu vermeiden, beinhaltet unsere Regressionsgleichung Interaktionsterme zwischen den drei Immigrationsdummies und unserem Proxy für den familiären Hintergrund. Unsere geschätzten EFH repräsentieren daher nur den Einfluss des familiären Hintergrundes für diejenigen Schüler, die selbst und deren Eltern ebenfalls in dem betreffenden Land geboren wurden.⁷

Wie im vorherigen Abschnitt bereits erläutert, wird dem besonderen Stichprobendesign auch hier durch Cluster-robuste und gewichtete Kleinst-Quadrat-Schätzungen Rechnung getragen. Tabelle 2 berichtet die Schätzer β_c der Stärke des Einflusses des familiären Hintergrundes (EFH) in den 54 Ländern. Die Ergebnisse belegen wiederum, dass die Schülerleistungen in jedem der betrachteten Länder signifikant mit dem familiären Hintergrund variieren.⁸ Der Wert der geschätzten Koeffizienten gibt an, um wie viele Punkte sich die Testergebnisse der Schüler in dem jeweiligen Land zwischen den fünf Kategorien der Bücher-Variablen durchschnittlich unterscheiden. Da die Testpunktzahl der Schüler wiederum so standardisiert wurde, dass die internationale Standardabweichung 100 beträgt, können die Koeffizienten als prozentualer Anteil einer internationalen Standardabweichung interpretiert werden, um den die Testleistung steigt, wenn sich die Anzahl der zu Hause verfügbaren Bücher

⁷ Für eine Analyse der Leistungseffekte von Immigrationspolitik vgl. Entorf und Minoiu (2005).

⁸ Die einzige Ausnahme ist Kuwait, dessen Schätzer nur auf dem 14-Prozent-Niveau statistische Signifikanz erreicht. Der Schätzer für Kolumbien ist auf dem 5-Prozent-Niveau statistisch signifikant, während die Schätzer für alle anderen Länder auf dem 1-Prozent-Niveau statistisch signifikant sind.

um eine Kategorie erhöht.⁹ Die geschätzte Stärke des EFH für Deutschland von 25,6 ist demnach so zu interpretieren, dass ein Unterschied des familiären Hintergrundes, der zwischen Familien mit einem Regal voll Büchern und zwei Regalen voll Büchern besteht, sich für die betreffenden Kinder in der Testleistung mit einem Unterschied von 25,6 Punkten bemerkbar macht.

Wie aus Abbildung 1 hervorgeht, die die geschätzten EFH der OECD-Länder nach absteigender Größe anordnet, liegt Deutschland damit an vierter Stelle der ungleichsten OECD-Länder. Eine noch stärkere Chancenungleichheit weisen nur England (28,8), Schottland (27,0) und Ungarn (25,8) auf. Am unteren Ende der Skala befinden sich als Länder mit den geringsten geschätzten Leistungsunterschieden zwischen Schülern mit unterschiedlichem Familienhintergrund Frankreich (8,3), Kanada (9,8), Portugal (10,4) und das flämische Schulsystem in Belgien (11,0).

C. Bestimmungsgründe der Chancengleichheit im Schulsystem

Die Ergebnisse des vorigen Abschnitts belegen eine ausgeprägte Variation in dem innerhalb der einzelnen Schulsysteme erreichten Ausmaß an Chancengleichheit. Im Folgenden untersuchen wir, ob und wie diese Variation durch unterschiedliche Merkmale der Schulsysteme zu erklären sind. Die Ergebnisse tragen zur Beantwortung der Frage bei, warum die Schulsysteme mancher Länder eine höhere Chancengleichheit erzielen als andere.

I. Identifikation durch Interaktionseffekte auf Schülerebene

a. Merkmale von Schulsystemen auf Länderebene

Es ist zu erwarten, dass verschiedene Merkmale der Schulsysteme mit den geschätzten EFH in Zusammenhang stehen. Ein Systemmerkmal, das zwischen den teilnehmenden Ländern erheblich variiert, ist das Alter der Schüler zum Zeitpunkt der ersten schulischen Selektion, bei der Schüler nach ihrer individuellen Leistungsfähigkeit in unterschiedliche Schulformen oder Laufbahnen aufgeteilt werden. In der Mehrheit der Länder findet eine Selektion erst mit 14 Jahren oder später statt, so dass viele Schüler zum Zeitpunkt des Tests (bei dem die meisten Schüler 13-14 Jahre alt sind) noch keine Selektionsstufe durchlaufen haben. Im Gegensatz dazu findet in manchen Ländern, wie auch in den meisten deutschen Bundesländern, bereits mit 10 Jahren die erste schulische Selektion statt. Es lässt sich vermuten, dass die schulische

⁹ Aufgrund mangelnder Daten kontrollieren die für Frankreich und Japan geschätzten EFH nicht für den Immigrationsstatus der Eltern, und in Frankreich zusätzlich nicht für den Immigrationsstatus der Schüler.

Leistung der Schüler um so stärker von ihrem familiären Umfeld beeinflusst wird, je jünger sie sind. Deshalb ist anzunehmen, dass eine frühe Selektion vor allem Schülern aus sozioökonomisch schwachen Familien schadet und dadurch die Gleichheit der Bildungschancen reduziert.¹⁰

Eine graphische Gegenüberstellung der EFH-Schätzer mit dem Alter der Schüler bei der ersten schulischen Selektion für die OECD-Länder zeigt, dass diese Einschätzung eine gewisse Berechtigung hat (Abbildung 1). Zwar ist die Aussagekraft dieser zweidimensionalen Abbildung beschränkt (die Ergebnisse einer multivariaten Analyse werden weiter unten berichtet). Doch die EFH-Schätzer derjenigen Länder, die ihre Schüler bereits mit 10 Jahren unterschiedlichen Schulformen zuweisen, liegen ausschließlich in der oberen Hälfte der Verteilung, während sich im unteren Drittel der Verteilung ausschließlich solche Schulsysteme finden, die ihre Schüler erst mit 14 Jahren oder später selektieren.

Weiterhin ist zu vermuten, dass die Gleichheit der Bildungschancen umso größer ist, je früher Kinder – speziell aus Familien mit niedrigem sozioökonomischen Status – mit formaler Bildung in Kontakt kommen. Bevor Kinder zum ersten Mal Bildungsinstitutionen besuchen, ist ihr „Bildungsniveau“ vor allem von ihren Familien bestimmt. Sobald sie jedoch formale Bildungseinrichtungen besuchen, können diese einen zusätzlichen und ausgleichenden Effekt auf die Lernfortschritte von Kindern ausüben. Deshalb ist zu erwarten, dass die Chancengleichheit für Kinder mit unterschiedlichem Familienhintergrund umso höher ist, je früher dieser ausgleichende Effekt einsetzt. Diese Hypothese lässt sich anhand von Informationen über die Besuchsquote im Vorschulbereich (Kindergarten), die Dauer der vorschulischen Erziehung und das Alter, an dem die Schulpflicht einsetzt, überprüfen. Es ist jedoch zu vermuten, dass der Zusammenhang zwischen EFH und der Besuchsquote im Vorschulbereich eine nicht-lineare Form besitzt, weil es anfänglich vor allem die Kinder aus wohlhabenden Familien sind, die den Kindergarten und andere vorschulische Einrichtungen besuchen. Erst wenn ein substantieller Anteil der Kinder in der entsprechenden Altersstufe, und damit auch Kinder von weniger begünstigten Familien, vorschulische Bildungseinrichtungen besucht, kann die vorschulische Erziehung einen ausgleichenden Effekt erzielen.

Ebenso wie die in formalen Bildungsinstitutionen verbrachten Jahre könnte auch die Länge des Schultages die Gleichheit der Bildungschancen beeinflussen. Es ließe sich vermuten, dass die Stärke des Familieneinflusses umso geringer ist, je mehr Zeit die Kinder in der Schule verbringen. Unter der Annahme, dass in Schulsystemen mit Ganztagsunterricht die Betreuung

¹⁰ Für eine modelltheoretische Begründung der hier angestellten Hypothesen vgl. Schütz et al. (2005).

intensiver ausfällt und sich beispielsweise auch auf die Hausaufgabenbetreuung erstreckt, könnten sie die Benachteiligung von Kindern aus sozioökonomisch schwachen Familien verringern.

Weitere berücksichtigte mögliche Einflussfaktoren sind das Niveau der staatlichen Bildungsausgaben pro Schüler, die durchschnittliche Testleistung in dem jeweiligen Land, das Niveau der wirtschaftlichen Entwicklung sowie der Anteil des privaten Sektors an den Bildungsausgaben und den Einschulungsquoten. Daten zu allen diesen Merkmalen stammen überwiegend aus statistischen Jahrbüchern und Datensammlungen von internationalen Organisationen wie der UNESCO und der OECD sowie aus gezielten länderspezifischen Anfragen. Besondere Aufmerksamkeit wurde darauf gerichtet, die jeweilige Information in dem für die getesteten Schülerpopulationen relevanten Zeitrahmen zu erfassen.¹¹

b. Empirische Spezifikationen

Um die Zusammenhänge zwischen Chancenungleichheit und den Schulsystemmerkmalen empirisch zu analysieren, sind zwei Vorgehensweisen denkbar. Einerseits kann man die geschätzten EFH direkt als zu erklärende Variable auf die Merkmale der Schulsysteme auf Länderebene regressieren. Andererseits kann der Zusammenhang zwischen Chancengleichheit und Schulsystem auch über Interaktionseffekte auf Schülerebene identifiziert werden.

In der Spezifikation auf Länderebene werden die geschätzten EFH (β_c aus Gleichung (2)) als abhängige Variable auf die potentiellen erklärenden Variablen Z_c der Schulsysteme regressiert:

$$\beta_c = \alpha + \theta Z_c + \mu_c . \quad (3)$$

Bei dieser Schätzgleichung ist die abhängige Variable selbst das Ergebnis einer Regressions-schätzung und keine präzise Beobachtung. In der Schätzung müssen daher die unterschiedlichen Standardfehler, mit denen die EFH geschätzt wurden, berücksichtigt werden. Dazu nutzen wir das von Anderson (1993) vorgeschlagene Verfahren, das jene Beobachtungen weniger stark gewichtet, welche relativ ungenau geschätzt wurden. Dabei werden zunächst die quadrierten Residuen einer Kleinst-Quadrate-Schätzung der Gleichung (3) berechnet und in einer zweiten Regression auf die geschätzten Varianzen der EFH, die quadrierten Varianzen und die Varianzen hoch drei regressiert. Diese Regression gibt an, in welchem Ausmaß die

¹¹ Definitionen, Quellenangaben und deskriptive Statistiken für die in den Schätzgleichungen verwendeten Variablen finden sich in Schütz et al. (2005).

quadratierten Residuen der ersten Regression durch die Varianzen der EFH-Schätzer erklärt werden können. Im letzten Schritt werden die invertierten vorhergesagten Werte dieser Regression als Gewichtungsfaktoren in einer gewichteten Kleinst-Quadrate-Schätzung der Gleichung (3) verwendet. Dadurch werden solche Beobachtungen weniger stark gewichtet, die relativ unpräzise geschätzt wurden.

Die zweite, alternative Spezifikation identifiziert auf Schülerebene die Zusammenhänge zwischen Chancenungleichheit und Merkmalen des Schulsystems über Interaktionseffekte zwischen dem familiären Hintergrund und den Schulsystemvariablen. Die abhängige Variable ist nun wie in Gleichung (2) die Testpunktzahl des jeweiligen Schülers, nur dass die Schätzung nun die Observationen aller Teilnehmerländer gemeinsam berücksichtigt:

$$T_{isc} = \alpha + \beta B_{isc} + (B_{isc} Z_c) \eta + F_{2,isc} \gamma + C_c \delta_1 + (C_c F_{2,isc}) \delta_2 + \varepsilon_{isc} , \quad (4)$$

wobei B_{isc} unsere Proxyvariable für den familiären Hintergrund und Z_c der Vektor der auf Länderebene gemessenen systemischen Variablen ist. Der Vektor $F_{2,isc}$ enthält dieselben Kontrollvariablen auf Schülerebene wie Gleichung (2). Der Vektor C_c enthält ein komplettes Set von Länderdummies und kontrolliert damit für fixe Effekte auf Länderebene. Dadurch berücksichtigt die Gleichung mögliche Niveauheterogenität zwischen den Ländern, und über die Interaktionsterme $C_c F_{2,isc}$ mögliche Heterogenität in den Effekten der individuellen Kontrollvariablen. Der zu schätzende Parametervektor η beschreibt den gesuchten Zusammenhang zwischen den Merkmalen des Schulsystems und der innerhalb des jeweiligen Schulsystems vorgefundenen Chancenungleichheit.

Bei der Schätzung des Fehlerterms muss diesmal berücksichtigt werden, dass die Schulsystemmerkmale, die zusammen mit der Variable für den familiären Hintergrund den für uns interessanten Effekt identifizieren, auf der Ebene der Länder gemessen werden und auch nur zwischen den Ländern variieren. Die primäre Erhebungseinheit in Gleichung (4) ist deshalb das Land. Unsere Regressionen lassen daher Interdependenzen in der Varianz-Kovarianz-Matrix der Fehlerterme innerhalb eines Landes zu und erfordern lediglich die Unabhängigkeit von Beobachtungen zwischen den Ländern. Die Standardfehler werden dadurch so berechnet, als gäbe es nur so viele Beobachtungen, wie es Länder gibt. Zusätzlich zu der Gewichtung der Schüler innerhalb eines Landes nach ihrer Erhebungswahrscheinlichkeit werden die Beobachtungen nun auch so gewichtet, dass jedes Land innerhalb der Regression das gleiche Gewicht erhält.

c. Ergebnisse

Tabelle 3 berichtet in den Spalten (1) und (2) die Ergebnisse der Spezifikation auf Länderebene und in den Spalten (3) bis (5) die Ergebnisse der Interaktionsspezifikation auf Schülerebene. Die grundlegenden Ergebnisse beider Spezifikationen sind identisch. Da jedoch der Spezifikation auf Schülerebene weniger restriktive Annahme zugrunde liegen, werden wir uns im Folgenden auf die Diskussion der Spalten (3) bis (5) konzentrieren.¹²

Das Basismodell in Spalte (3) beinhaltet als erklärende Variablen Interaktionsterme zwischen der Variable für den familiären Hintergrund und folgenden Schulsystemmerkmalen: Alter der Schüler bei der ersten schulischen Selektion, Besuchsquote im Vorschulbereich und deren Quadrat sowie Dauer der Vorschulprogramme. In Übereinstimmung mit dem erwarteten Effekt sinkt die Chancenungleichheit (der EFH) mit zunehmendem Alter der Schüler bei der ersten schulischen Selektion. Je früher Schüler also in unterschiedliche Schulformen selektiert werden, umso höher ist die Chancenungleichheit. Dieser Zusammenhang ist statistisch hoch signifikant.

Der Zusammenhang zwischen EFH und den Besuchsquoten im Vorschulbereich folgt einem umgekehrt U-förmigen Verlauf und ist ebenfalls statistisch hoch signifikant. Solange nur ein geringer Anteil der entsprechenden Schülerpopulation vorschulische Einrichtungen besucht, hängen die Besuchsquoten im Vorschulbereich positiv mit dem EFH-Schätzer zusammen. Ein solcher Zusammenhang ist als Folge einer nicht zufälligen Selektion von Schülern mit „gutem“ familiären Hintergrund in Vorschulprogramme durchaus denkbar. Erst wenn ein substantieller Anteil der Kinder vorschulische Einrichtungen besucht, so dass auch Kinder aus Familien mit „schlechterem“ sozioökonomischen Hintergrund darunter sind, beginnen sich höhere Besuchsquoten im Vorschulbereich negativ auf den geschätzten Einfluss des Familienhintergrundes auszuwirken. Auf der Grundlage der geschätzten Parameter (Spalte (3)) ergibt sich, dass bis zu einer Besuchsquote von 61 Prozent die Bildungschancen mit steigenden Besuchsquoten ungleicher werden. Oberhalb dieses Schwellenwertes sinkt die Chancenungleichheit mit weiterhin steigenden Besuchsquoten.

Auch die Dauer des Vorschulprogramms zeigt den vermuteten Effekt. Die Stärke des Einflusses des familiären Hintergrundes auf die Schülerleistung sinkt in statistisch hoch signifikanter Weise mit zunehmender offizieller Dauer der Vorschulprogramme. Zusammenge-

¹² Taiwan konnte in allen Regressionen der Tabelle 3 nicht berücksichtigt werden, da international vergleichbare Angaben zu den meisten Schulsystemvariablen fehlen.

nommen erklären diese drei Merkmale der Schulsysteme 40 Prozent der gesamten Variation der EFH-Schätzer (Spalte (1)).

Abbildung 2 veranschaulicht die mit Hilfe der Modellspezifikation von Spalte (3) geschätzten Interaktionseffekte. Dazu stellt die Graphik den Einfluss dieser drei Schulsystemmerkmale den EFH-Schätzern ausgewählter Länder gegenüber. Jedes Jahr, um das die Selektion von Schülern in verschiedene Schulformen aufgeschoben wird, verringert die geschätzten EFH um etwas mehr als eine Einheit, bzw. um etwas mehr als 1 Prozent einer internationalen Standardabweichung der Testpunktzahl. Ein Unterschied im Alter der Schüler zum Zeitpunkt der ersten schulischen Selektion von vier Jahren (im Alter von 14 Jahren wie in den meisten Ländern statt im Alter von 10 Jahren wie in Deutschland) ist also verbunden mit einem Unterschied in den geschätzten EFH von 4,3 Punkten. Dies entspricht ungefähr einem Viertel des gesamten durchschnittlichen EFH der OECD-Länder (17,7). Ähnlich führt eine Erhöhung der Besuchsquote im Vorschulbereich von 60 Prozent (z.B. in der Schweiz oder den USA) auf 100 Prozent (z.B. in Dänemark oder den Niederlanden) zu einer Verringerung der EFH um 4,4 Punkte. Eine Erhöhung der Dauer der Vorschulerziehung um 1 Jahr geht mit einer Verringerung der EFH um 1,3 Punkte einher.

In Spalte (4) der Tabelle 3 fügen wir der Modellspezifikation drei weitere Variablen hinzu. Weder die Bildungsausgaben pro Schüler noch das Bruttonationaleinkommen pro Kopf weisen jedoch einen statistisch signifikanten Interaktionseffekt mit dem familiären Hintergrund auf. Der Interaktionseffekt zwischen der durchschnittlichen Testpunktzahl eines Landes und dem familiären Hintergrund ist dagegen statistisch signifikant und positiv. Wenn man jedoch die Regression von Spalte (4) für die homogenere Gruppe der OECD-Staaten durchführt, ist der Effekt negativ und gleichzeitig statistisch nicht mehr signifikant (vgl. Schütz et al. 2005 für Details). Innerhalb der OECD-Staaten gibt es also keinen Beleg für einen Zielkonflikt zwischen Effizienz im Sinne einer hohen durchschnittlichen Testpunktzahl und Gleichheit im Sinne eines geringen Einflusses des familiären Hintergrundes auf die Schülerleistung. Bei den Ergebnissen der schon im Basismodell in Spalte (3) betrachteten Variablen ergibt sich durch die zusätzliche Berücksichtigung der drei weiteren erklärenden Variablen keine qualitative Veränderung.

Das Modell in Spalte (5) berücksichtigt noch zwei weitere Variablen: den Anteil der aus privaten Quellen stammenden Bildungsausgaben an den gesamten Bildungsausgaben und den Anteil der Schüler, der in Privatschulen unterrichtet wird. In dieser Spezifikation fällt die Anzahl der beobachteten Schulsysteme aufgrund mangelnder international vergleichbarer Daten

auf 27. Die geschätzte Stärke des Einflusses des familiären Hintergrundes steigt statistisch signifikant mit dem Anteil der privaten Bildungsausgaben, während sie mit dem Anteil der Schüler an Privatschulen statistisch signifikant fällt. Bildungssysteme, die einen höheren Anteil privater Bildungsausgaben aufweisen, weisen eine höhere Ungleichheit auf, während ein höherer Anteil von Privatschulen die Ungleichheit zu reduzieren scheint.

Um auszuschließen, dass Variationen zwischen sehr heterogenen Ländern für die hier dargestellten Ergebnisse verantwortlich sind, wurden die Regressionsgleichungen der Tabelle 3 auch für die kleinere und homogenere Gruppe der OECD-Staaten geschätzt. Dabei erweisen sich alle zuvor berichteten Effekte als robust gegenüber dieser Einschränkung der Stichprobe (vgl. Schütz et al. 2005). Als weiterer Robustheitstest wurden Frankreich und Japan aus der Stichprobe entfernt, weil beide Länder keine Daten zu dem Immigrationsstatus der Eltern erhoben haben, und Frankreich darüber hinaus auch keine Daten zum Immigrationsstatus des Schülers. Wiederum hat die veränderte Stichprobe keinen Einfluss auf die Ergebnisse.

Als weitere potentielle erklärende Variablen auf der Ebene der Schulsysteme wurden Ganztags- bzw. Halbtagsunterricht und der Beginn der Schulpflicht betrachtet. Informationen über die Länge des Schultages lagen für 34 Schulsysteme vor. Daraus wurde eine Dummyvariable gebildet, die den Wert 1 annimmt, wenn die Schule bis zum Nachmittag dauert, und 0, wenn die Schule gegen Mittag zu Ende ist. In keiner der betrachteten Spezifikationen zeigt der Dummy einen signifikanten Zusammenhang mit den EFH. Ein Datenproblem besteht in diesem Ergebnis darin, dass die verfügbaren Daten keine Aussage darüber zulassen, ob nachmittags Unterricht stattfindet oder ob für die Schüler nur eine optionale Betreuung angeboten wird. Auch die Ergebnisse zum Zusammenhang zwischen den EFH und dem Alter zu Beginn der Schulpflicht sind nicht eindeutig. Der Grund dafür lässt sich darin vermuten, dass das Alter zu Beginn der Schulpflicht zwischen den Ländern nicht stark variiert (zwischen 5 und 7 Jahren) und dass von ihm kein zusätzlicher Effekt mehr ausgeht, nachdem bereits für die Besuchsquote und die Dauer des Vorschulbereichs kontrolliert worden ist.

II. Identifikation durch Differenzen in Differenzen

Ein verbleibendes Problem jeglicher internationaler Vergleiche, und damit auch der gerade präsentierten, besteht darin, dass sie sich möglicherweise einer unbeobachteten Heterogenität von Staaten gegenüber sehen. Es ist vorstellbar, dass viele weitere, oftmals unbeobachtete Faktoren die beobachteten Leistungen beeinflussen, und nur unter der (im Bisherigen impliziten) Annahme, dass solche unbeobachtete Heterogenität zwischen den Ländern unabhängig

ist vom Einfluss der betrachteten Variablen auf den EFH, spiegeln die berichteten Ergebnisse tatsächliche Effekte wider. Bei systematischer Heterogenität könnte es dagegen zu Verzerrungen kommen. Für einen Einflussfaktor, der sich in den bisherigen Untersuchungen als relevant erwiesen hat, nämlich die frühe Selektion in einem mehrgliedrigen Schulsystem, schildern wir im Folgenden die Ergebnisse einer zusätzlichen Spezifikation, die möglicher unbeobachteter Heterogenität Rechnung trägt. Das bisherige Ergebnis, dass frühe Selektion die Chancenungleichheit erhöht, erweist sich in dieser Spezifikation als robust.

a. Die Identifikationsidee

Die Kernidee der alternativen Spezifikation besteht darin, dass die Mehrgliedrigkeit überall erst nach der Grundschule eingeführt wird. Damit ergibt sich die Möglichkeit, schon vor der Mehrgliedrigkeit bestehende Unterschiede herauszurechnen, wenn man zusätzlich zu den bisher betrachteten Mittelstufenstudien auch Grundschulstudien berücksichtigt. Durch Gegenüberstellung von Tests am Ende der Grundschule und Tests gegen Ende der Mittelstufe (Sekundarstufe 1) lässt sich dann im Rahmen einer internationalen Differenzen-in-Differenzen-Schätzung (DiD) der Einfluss der Mehrgliedrigkeit λ ermitteln. Dabei wird die Leistungsveränderung von der Grund- zur Sekundarschule zwischen ein- und mehrgliedrigen Schulsystemen verglichen:

$$\begin{aligned} \lambda^{DiD} &= \left(\bar{Y}_{Treatmentgruppe}^{nach} - \bar{Y}_{Treatmentgruppe}^{vor} \right) - \left(\bar{Y}_{Kontrollgruppe}^{nach} - \bar{Y}_{Kontrollgruppe}^{vor} \right) \\ &= \Delta \bar{Y}_{Treatmentgruppe} - \Delta \bar{Y}_{Kontrollgruppe} \end{aligned} \quad (5)$$

wobei Y für das jeweils betrachtete Ergebnis, z.B. die Streuung oder das Niveau der Schülerleistungen, steht.

Wie im klassischen DiD-Modell erfolgt die Identifikation des Effekts hier also über den Unterschied zwischen einer „Treatment“-Gruppe (Systeme, in denen Mehrgliedrigkeit vorliegt) und einer Kontrollgruppe (bis ans Ende der Mittelstufe eingliedrige Systeme) in der Differenz des Ergebnisses vor und nach Einführung der Mehrgliedrigkeit. Dabei besteht aber im Gegensatz zum klassischen DiD-Modell der Unterschied zwischen „vorher“ und „nachher“ nicht in der zeitlichen Dimension, sondern im Vergleich über die Jahrgangsstufen. In der Regressionsanalyse wird dabei das Grundschulergebnis jedes Landes als Kontrollvariable für sein eigenes Ergebnis in der Sekundarschule genutzt. Damit wird jegliche Heterogenität zwischen Ländern, die schon am Ende der Grundschule besteht, aus dem Effekt der Mehrgliedrigkeit herausgerechnet. Insofern – und das ist die verbleibende identifizierende Annahme –

keine weiteren Einflüsse die zwischen Grundschule und Mittelstufe auftretende Ergebnisveränderung in einer in systematischem Zusammenhang mit der Mehrgliedrigkeit stehenden Weise beeinflussen, kann der DiD-Schätzer als kausaler Einfluss der Mehrgliedrigkeit interpretiert werden.¹³

Als zentrales Maß des Ergebnisses Y betrachten wir im Folgenden die nationale Standardabweichung der Testpunktzahl als Maß der Streuung der Schülerleistungen in einem Land. Streuungsmaße geben wieder, inwiefern es einem Schulsystem gelingt, alle Schüler auf den gleichen Wissensstand zu bringen und gleichermaßen zu fördern und können insofern als Maße der innerhalb des Schulsystems verwirklichten Chancengleichheit betrachtet werden. Da die Gesamtstreuung der Leistungen aber nicht nur von der Größe des Einflusses des familiären Hintergrundes (wie etwa des zuvor berechneten EFH) abhängt, sondern auch von der Streuung im familiären Hintergrund selbst, ist es allerdings nur ein beschränktes Maß der durch das Schulsystem bedingten Chancenungleichheit. Allerdings sollten in der vorliegenden Analyse internationale Unterschiede im durchschnittlichen familiären Hintergrund aber aufgrund der verwendeten DiD-Schätzung keine Rolle spielen.¹⁴ Über das Ungleichheitsmaß hinaus werden wir auch kurz den Einfluss der Mehrgliedrigkeit auf das Leistungsniveau eines Landes betrachten, um die Möglichkeit eines Zielkonflikts zwischen Gleichheit und Leistungsniveau beleuchten zu können.

b. Ergebnisse

Zur Implementierung der vorgeschlagenen DiD-Strategie bedarf es zusammenfügbarer Leistungsdaten aus Grund- und Sekundarschule. Dementsprechend betrachten wir insgesamt acht Paare internationaler Vergleichstests (vgl. Hanushek und Wößmann 2005 für Details zu den Daten). Beim ersten Paar ist die Sekundarschulstudie der PISA-2003-Test, der mit der IGLU-Grundschulstudie zusammengebracht wird. Beide Studien testen Leseleistungen, PISA unter 15-Jährigen, IGLU unter Viertklässlern. Insgesamt sind es 18 Länder, die an beiden Tests teilgenommen haben. Der zentrale Befund in Bezug auf den Einfluss von Mehrgliedrigkeit auf Chancenungleichheit wird in Abbildung 3 graphisch dargestellt. Auf der rechten Seite ist die Standardabweichung der PISA-Leistungen als Maß der Leistungsstreuung eines jeden

¹³ Ein einfaches analytisches Modell, das die vorliegende Identifikation näher herleitet, sowie eine umfassendere Diskussion der Identifikationsstrategie finden sich in Hanushek und Wößmann (2005).

¹⁴ Tatsächlich besteht eine signifikante Korrelation zwischen der Standardabweichung der TIMSS-Testpunktzahl in einem Schulsystem und dem zuvor verwendeten Schätzer für die Stärke des Einflusses des familiären Hintergrundes (EFH) in dem gleichen Schulsystem; der Korrelationskoeffizient beträgt 0,64. Somit erweisen sich familiäre Hintergrundeffekte als ein wichtiger Erklärungsfaktor der gesamten Leistungsstreuung.

Landes, relativ zum Durchschnitt der 18 Länder, abgetragen. In diesem reinen Querschnitt hängt die Leistungsstreuung nicht systematisch mit dem Vorliegen von Mehrgliedrigkeit zusammen. So befinden sich etwa sowohl unter den sechs „ungleichsten“ als auch unter den sechs „gleichsten“ Ländern jeweils drei Länder, in denen die 15-Jährigen in verschiedenen Schularten unterrichtet werden, und drei Länder, in denen die 15-Jährigen (noch) nicht in verschiedenen Schultypen selektiert wurden. Aber wie bereits erläutert könnte dieser Befund durch unbeobachtete Heterogenität zwischen den betrachteten Ländern verzerrt sein.

Dementsprechend besteht die Grundidee der DiD-Strategie darin, auf die *Veränderung* in der Streuung zwischen Grund- und Sekundarschulbereich zu schauen. Dazu bildet Abbildung 3 auf der linken Seite die Standardabweichung der IGLU-Grundschulleistungen ab. Die Steigung der Linien zwischen den beiden Seiten gibt nun an, wie sich die Leistungsstreuung in den verschiedenen Systemen zwischen Grund- und Sekundarschule verändert hat. Der zentrale Befund dieser Betrachtungsweise ist direkt augenfällig: Die Ungleichheit steigt systematisch zwischen Grund- und Sekundarschule in Ländern mit früher Mehrgliedrigkeit, und sie sinkt systematisch in Ländern ohne Mehrgliedrigkeit. Die Verbindungslinien für nahezu alle Länder, die im Alter von 15 Jahren mehrere Schultypen aufweisen, streben nach oben, wohingegen sie für nahezu alle Länder mit noch eingliedrigem Schulsystem nach unten weisen. Dabei erweist sich Deutschland als das Land, in dem die Ungleichheit zwischen der Grundschule und dem Ende der Mittelstufe am stärksten von allen betrachteten Ländern ansteigt. Wie aus der Abbildung leicht ersichtlich, weist Deutschland am Ende der Grundschule eine international durchaus noch unterdurchschnittliche Ungleichheit auf, welche erst im Verlaufe der mehrgliedrigen Mittelstufe bis zum Alter von 15 Jahren zur größten Ungleichheit in der betrachteten Länderstichprobe anwächst.

Systematisch lässt sich dieser Befund im Rahmen einer Regressionsanalyse zeigen. Die Ergebnisse werden in Spalte (1) von Tabelle 4 berichtet, wozu die Testleistungen in den beiden Tests jeweils standardisiert wurden (vgl. Hanushek und Wößmann 2005). Im oberen Teil der als System geschätzten Zusammenhänge finden sich die Ergebnisse für den Einfluss der frühen Mehrgliedrigkeit auf die Leistungsstreuung. Nachdem für die schon in der Grundschule bestehende Streuung, welche einen signifikant positiven Einfluss auf die Streuung im Sekundarschulbereich hat, kontrolliert wird, besteht ein statistisch und quantitativ signifikanter positiver Zusammenhang zwischen früher Mehrgliedrigkeit und Leistungsstreuung im Sekundarschulbereich. Die Effektgröße von 0,25 zeigt, dass die Mehrgliedrigkeit quantitativ ein Viertel des gesamten Unterschieds in der Ungleichheit zwischen dem ungleichsten

(Deutschland mit einer standardisierten Streuung von 4,5) und dem gleichsten Land (Hongkong mit 3,5) ausmacht.

Im unteren Teil der Tabelle 4 wird der Einfluss der Mehrgliedrigkeit auf das Leistungsniveau berichtet. Zum einen besteht ein starker Zusammenhang zwischen dem Leistungsniveau der 15-Jährigen und dem bereits in der Grundschule observierbaren Leistungsniveau. Unter Berücksichtigung dieses Effektes ergibt sich in Spalte (1) dann ein statistisch signifikanter *negativer* Effekt der frühen Mehrgliedrigkeit auf das erzielte durchschnittliche Leistungsniveau. Das bedeutet, dass mehrgliedrige Systeme mit der höheren Ungleichheit nicht etwa ein höheres Durchschnittsniveau erzielen, sondern im Gegenteil sogar ein niedrigeres.¹⁵

Die Spalten (2) bis (8) berichten Schätzungen derselben DiD-Spezifikation für sieben weitere Paare internationaler Vergleichstests. Dabei legen die Spezifikation (2) bis (6) Wert auf eine möglichst zeitnahe Verbindung von Primar- und Sekundarschultest, um mögliche Einflüsse anderer zeitvariater Bedingungen zu minimieren. Demgegenüber verfolgen die Spezifikationen (7) und (8) eine spezifische Schülerkohorte über die Zeit: Die Kohorte, die 1995 von TIMSS in der vierten Klasse getestet wurde, wurde vier Jahre später von TIMSS-Repeat getestet, als sie 1999 in der achten Klasse war. Der generelle Befund in Bezug auf den Zusammenhang von Mehrgliedrigkeit und Ungleichheit ist eindeutig: sieben der acht Schätzer weisen einen positiven Koeffizienten aus, von denen fünf statistisch signifikant sind. Der Befund in Bezug auf das Leistungsniveau ist weniger einheitlich: Während zwei weitere Paare einen statistisch signifikanten negativen Zusammenhang ausweisen, liegt auch ein statistisch signifikanter positiver Koeffizient vor.

Die letzte Spalte von Tabelle 4 weist das Ergebnis einer Schätzung aus, in der alle acht Paare internationaler Leistungstests gemeinsam betrachtet werden. In dieser gepoolten Schätzung findet sich ein klarer positiver Effekt der frühen Mehrgliedrigkeit auf die Leistungsstreuung und ein weniger eindeutiger, aber statistisch noch signifikanter negativer Effekt auf das durchschnittliche Leistungsniveau. Die Ergebnisse wurden auch mehreren Robustheitstests unterworfen (vgl. Hanushek und Wößmann 2005 für Details). So ergeben sich qualitativ unveränderte Ergebnisse, wenn anstelle der Standardabweichung alternative Streuungsmaße wie der Leistungsunterschied zwischen dem 75. und 25. Perzentil eines Landes oder zwischen dem 95. und 5. Perzentil betrachtet werden. Die Ergebnisse sind auch qualitativ unverändert,

¹⁵ Hanushek und Wößmann (2005) berichten auch Effekte über die gesamte Verteilung, aus denen hervorgeht, dass sich die Mehrgliedrigkeit nicht nur auf die leistungsschwächeren Schüler, sondern bis hoch in das 95. Perzentil der Leistungsstreuung negativ auswirkt.

wenn anstelle des hier betrachteten Mehrgliedrigkeits-Dummies (liegt zum Zeitpunkt des Mittelstufentests Mehrgliedrigkeit vor oder nicht) ein lineares Mehrgliedrigkeitsmaß (Alter der Schüler bei der ersten Selektion) betrachtet wird. Trotz der geringen Freiheitsgrade in den relativ kleinen Länderstichproben hat der Befund auch bei Berücksichtigung einzelner zusätzlicher Kontrollvariablen wie des wirtschaftlichen Entwicklungsniveaus oder der Bildungsausgaben je Schüler Bestand.

D. Schlussbemerkungen

Im vorliegenden Beitrag wurde die in Schulsystemen bestehende Chancenungleichheit mit Hilfe der Daten verschiedener internationaler Schülerleistungstests geschätzt. Dabei wurde Chancenungleichheit als die Stärke des Einflusses des familiären Hintergrundes auf die Testleistung gemessen. Alle 54 betrachteten Länder weisen einen statistisch signifikanten Zusammenhang zwischen Schülerleistungen und familiärem Hintergrund auf.¹⁶ Gleichwohl gibt es im internationalen Vergleich in der Größe des Effektes des familiären Hintergrundes starke Unterschiede zwischen Ländern. In Großbritannien und Deutschland zeigt sich eine relativ große Chancenungleichheit für Kinder unterschiedlicher sozialer Herkunft, während Frankreich und Kanada deutlich ausgeglichene Leistungen aufweisen. Dementsprechend scheint es prinzipiell möglich zu sein, dass das Schulsystem das Ausmaß der Ungleichheit reduziert, indem es allen Kindern unabhängig von ihrem sozioökonomischen Hintergrund möglichst gleiche Chancen bietet.

Die berichteten ökonometrischen Analysen haben gezeigt, dass insbesondere zwei Merkmale des Schulsystems – die frühe Mehrgliedrigkeit und das Ausmaß des Vorschulsystems – in signifikanter Weise mit der erzielten Chancengleichheit zusammenhängen. Schwächere Evidenz deutet auch darauf hin, dass der Anteil privater Schulfinanzierung in einem negativen und der Anteil privater Schulträgerschaft in einem positiven Zusammenhang mit der erzielten Chancengleichheit stehen. Demgegenüber ließ sich für andere Merkmale des Schulsystems wie das Ausgabenniveau, die Durchschnittsleistung, die Länge des Schultages und das Bruttonationaleinkommen kein signifikanter Zusammenhang mit der erzielten Chancengleichheit nachweisen. Dieses Ergebnis legt nahe, dass Forderungen nach vermehrter Förderung von Ganztagschulen nicht unbedingt bildungspolitisch zu begründen sind, was eine familienpolitische Begründung natürlich nicht ausschließt.

¹⁶ Einzige Einschränkung ist Kuwait mit einem statistischen Signifikanzniveau von 14 Prozent.

In Bezug auf das Ausmaß des Vorschulsystems weisen sowohl die Dauer des Vorschulsystems als auch eine möglichst hohe Besuchsquote einen positiven Zusammenhang mit der Chancengleichheit auf. Dementsprechend ist die diesbezügliche kürzliche Empfehlung des Sachverständigenrats zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2004, S. 3) im Jahresgutachten 2004/05, in dem er den „Ausbau eines kostenlosen, aber auch verpflichtenden Förder- und Betreuungsangebots .. im Vorschulbereich“ für geboten ansieht, durchaus zu begrüßen.

In Bezug auf die frühe Mehrgliedrigkeit zeigt sich, dass der familiäre Einfluss umso größer ist, je eher die Selektion in unterschiedliche Schultypen erfolgt. Das mehrgliedrige Schulsystem wird oftmals mit angeblichen positiven Niveaueffekten, insbesondere für leistungsstarke Schüler, verteidigt. Die vorgelegten Befunde legen aber nahe, dass eine frühe Selektion der Schüler in verschiedene Schultypen nicht nur die Chancenungleichheit erhöht, sondern auch das gesamte Leistungsniveau sogar eher senkt als erhöht. Damit ergibt sich in diesem Bereich eher kein Zielkonflikt zwischen Gleichheit und Effizienz in der Organisation des Schulsystems.

Literatur

- Anderson, Patricia M. (1993). Linear Adjustment Costs and Seasonal Labor Demand: Evidence from Retail Trade Firms. *Quarterly Journal of Economics* 108 (4): 1015-1042.
- Baumert, Jürgen, Gundel Schümer (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Jürgen Baumert et al. (Hrsg.). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*, S. 323-407. Opladen: Westdeutscher Verlag.
- Beaton, Albert E., Ina V.S. Mullis, Michael O. Martin, Eugenio J. Gonzalez, Dana L. Kelly, Teresa A. Smith (1996). *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.
- Deaton, Angus (1997). *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore: The Johns Hopkins University Press.
- 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)
- Fuchs, Thomas, Ludger Wößmann (2004). What Accounts for International Differences in Student Performance? A Re-examination using PISA Data. CESifo Working Paper 1235. München: CESifo.
- Hanushek, Eric A., Ludger Wößmann (2005). Does Early Tracking Affect Educational Inequality and Performance? Difference-in-Difference Evidence across Countries. NBER Working Paper 11124. Cambridge, MA: National Bureau of Economic Research.
- Martin, Michael O., Dana L. Kelly (Hrsg.) (1996). *Third International Mathematics and Science Study Technical Report, Volume I: Design and Development*. Chestnut Hill, MA: Boston College.
- Moulton, Brent R. (1986). Random Group Effects and the Precision of Regression Estimates. *Journal of Econometrics* 32 (3): 385-397.
- Mullis, Ina V.S., Michael O. Martin, Eugenio J. Gonzalez, Kelvin D. Gregory, Robert A. Garden, Kathleen M. O'Connor, Steven J. Chrostowski, Teresa A. Smith (2000). *TIMSS 1999 International Mathematics Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. Chestnut Hill, MA: Boston College.
- Nickell, Stephen (2004). Poverty and Worklessness in Britain. *Economic Journal* 114 (494): C1-C25.
- Noël, Sylvain, Patrice de Broucker (2001). Intergenerational Inequities: A Comparative Analysis of the Influence of Parents' Educational Background on Length of Schooling and Literacy Skills. In: Walo Hutmacher, Douglas Cochrane, Norberto Bottani (Hrsg.). *In Pursuit of Equity in Education*, S. 277-297. London: Kluwer Academic Publishers.
- Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2004). *Erfolge im Ausland – Herausforderungen im Inland: Jahresgutachten 2004/05*. (Verfügbar unter www.sachverstaendigenrat-wirtschaft.de)

- Schütz, Gabriela, Heinrich W. Ursprung, Ludger Wößmann (2005). Education Policy and Equality of Opportunity. CESifo Working Paper 1518. München: CESifo.
- White, Halbert (1984). *Asymptotic Theory for Econometricians*. Orlando: Academic Press.
- Wooldridge, Jeffrey M. (2001). Asymptotic Properties of Weighted *M*-Estimators for Standard Stratified Samples. *Econometric Theory* 17 (2): 451-470.
- Wößmann, Ludger (2002). How Central Exams Affect Educational Achievement: International Evidence from TIMSS and TIMSS-Repeat. PEPG Working Paper 02-10. Cambridge, MA: Harvard University, John F. Kennedy School of Government, Program for Education Policy and Governance.
- Wößmann, Ludger (2003). Schooling Resources, Educational Institutions and Student Performance: The International Evidence. *Oxford Bulletin of Economics and Statistics* 65 (2): 117-170.
- Wößmann, Ludger (2004). How Equal Are Educational Opportunities? Family Background and Student Achievement in Europe and the United States. CESifo Working Paper 1162. München: CESifo.

Abbildung 1: Maß der Chancengleichheit im Schulsystem (EFH) in OECD-Ländern

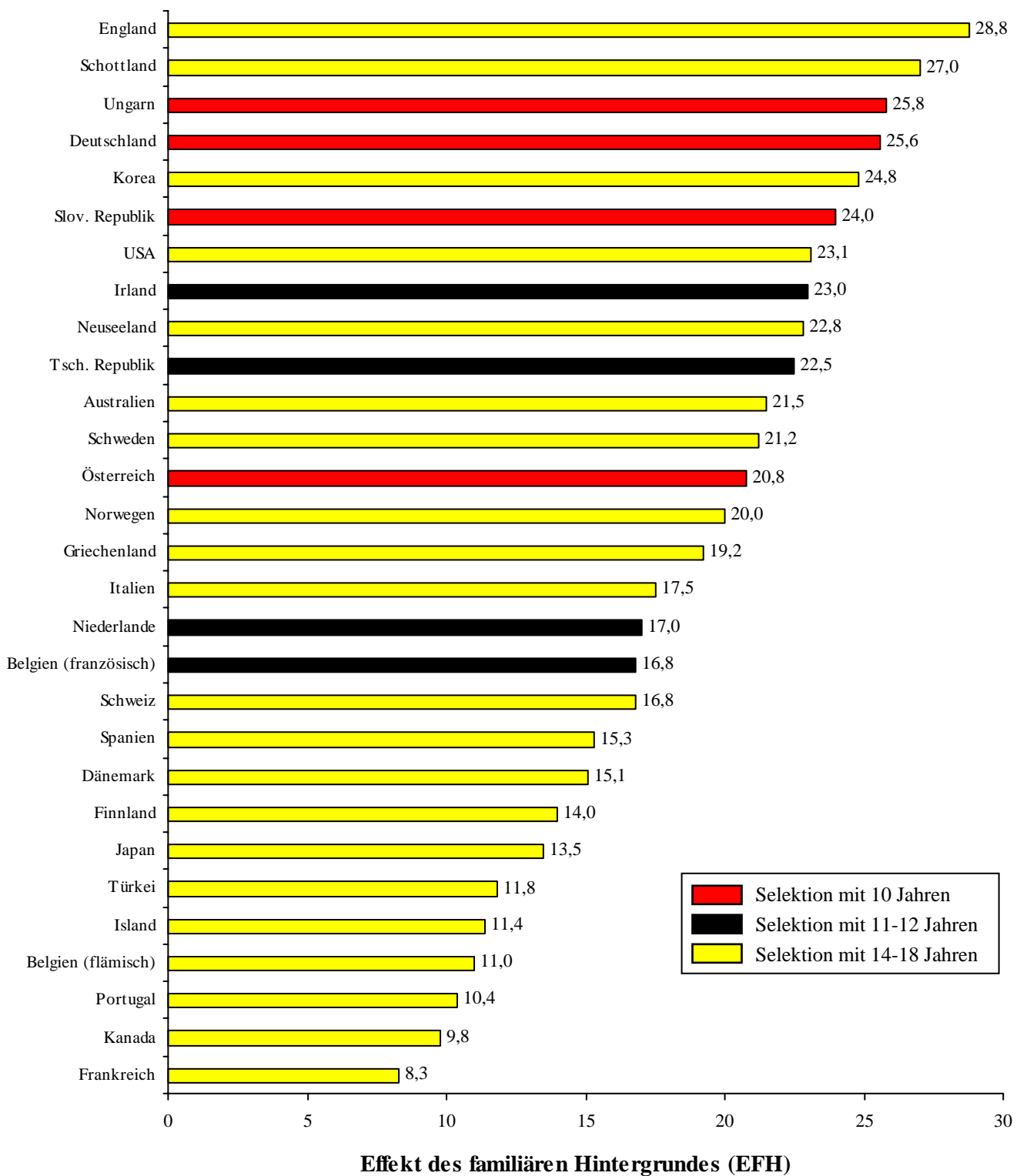


Abbildung 2: Bildungspolitik und Chancengleichheit

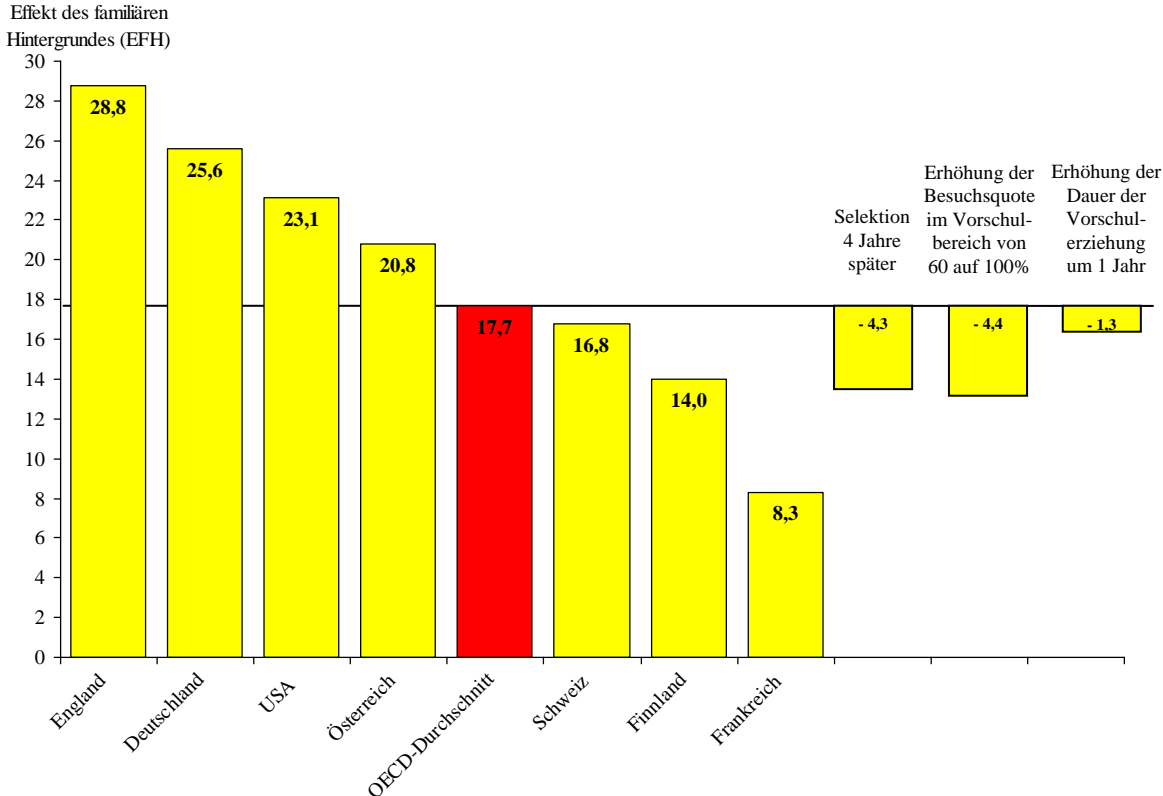


Abbildung 3: Bildungsungleichheit in Grundschule und Mittelstufe

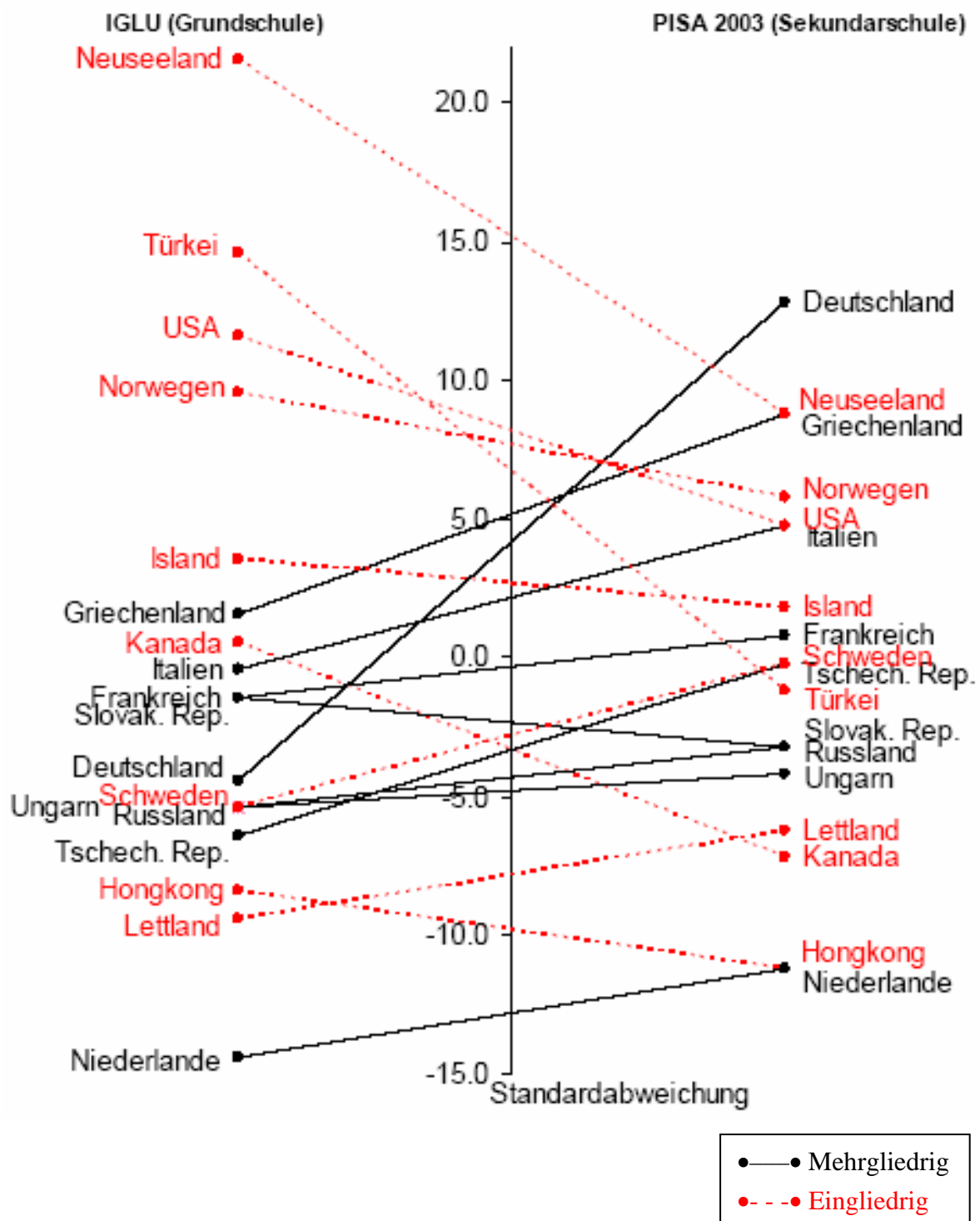


Tabelle 1: Familiäre Effekte in Europa und den USA

	Alle Länder	Interaktion mit Europadummy
Bildung der Eltern		
Etwas Sekundarbildung	16,56** (7,22)	-12,26* (7,34)
Sekundarabschluss	17,27** (7,33)	-4,39 (7,41)
Etwas über Sekundarbildung	30,83*** (7,01)	-8,73 (7,17)
Universitätsabschluss	52,01*** (8,02)	-20,08** (8,15)
Zu Hause verfügbare Bücher		
Ein Regalbrett (11-25)	9,25** (3,70)	7,60* (3,94)
Ein Regal (26-100)	34,82*** (3,58)	2,97 (3,84)
Zwei Regale (101-200)	53,42*** (4,20)	-0,39 (4,46)
Mehr als zwei Regale (>200)	62,41*** (4,67)	-0,87 (4,90)
Mit beiden Eltern lebend		
	15,40*** (2,82)	-9,69*** (2,96)
Im Land geboren		
	2,12 (4,69)	12,63*** (4,92)
Weiblich		
	-8,92*** (2,32)	0,00 (2,48)
Alter		
	-22,21*** (2,64)	3,47 (2,73)
Obere Klassenstufe		
	46,15*** (4,42)	4,96 (4,56)
Oberhalb der oberen Klassenstufe		
	111,60*** (3,29)	-
Lage der Schule		
Geographisch isoliert	-28,59*** (7,68)	25,77*** (8,56)
Nah am Stadtzentrum	-6,65 (6,30)	9,43 (6,50)
Beobachtungen (Schüler)	115.406	
Primäre Erhebungseinheiten (Schulen)	2.932	
R ²	0,262	

Kleinst-Quadrate-Schätzung (gewichtet mit Erhebungswahrscheinlichkeiten der Schüler), alle europäischen Länder und USA gepoolt. Abhängige Variable: TIMSS-Testpunkte in Mathematik. Schätzung kontrolliert für Länder- und Imputations-Dummies. Cluster-robuste Standardfehler (auf Schulebene) in Klammern. Signifikanzniveaus: *** 1%. – ** 5%. – * 10%.

Quelle: Wößmann (2004).

Tabelle 2: Basisdaten und geschätzter Einfluss des familiären Hintergrundes (EFH)

	EFH-Schätzung			Deskriptive Statistiken		
	EFH	Standardfehler	EFH Rang	TIMSS Testleistung (Durchschnitt)	TIMSS Standardabweichung	Zu Hause verfügbare Bücher ^a
Australien	21,45	(1,43)	19	541,4	87,1	3,9
Belgien (Flämisch)	10,95	(1,11)	45	558,2	73,3	3,0
Belgien (Französisch)	16,77	(1,77)	31	494,9	78,7	3,7
Bulgarien	23,32	(2,94)	9	529,2	83,1	3,6
Chile	21,95	(1,78)	18	423,4	78,2	2,6
Dänemark	15,08	(1,62)	37	486,4	78,3	3,8
Deutschland	25,57	(1,94)	5	517,5	89,3	3,5
England	28,81	(1,70)	1	530,0	87,3	3,6
Finnland	13,98	(1,51)	39	543,5	65,4	3,4
Frankreich	8,32	(1,44)	49	514,9	67,5	3,3
Griechenland	19,22	(1,63)	25	486,8	80,0	3,1
Hongkong	10,82	(1,28)	46	562,1	77,7	2,5
Indonesien	4,83	(1,81)	53	433,2	86,5	2,2
Iran	11,25	(0,93)	44	447,6	68,0	2,2
Irland	23,04	(1,98)	12	530,4	87,2	3,3
Island	11,42	(2,55)	43	486,3	70,5	4,0
Israel	18,84	(2,17)	26	501,9	94,1	3,4
Italien	17,51	(1,42)	29	502,0	80,8	3,1
Japan	13,50	(1,13)	40	583,5	80,4	3,1
Jordanien	17,92	(2,40)	28	455,1	94,8	2,6
Kanada	9,76	(0,95)	48	537,0	75,6	3,7
Kolumbien	7,55	(3,84)	50	390,5	62,4	2,4
Korea	24,75	(0,81)	6	584,7	85,5	3,3
Kuwait	2,49	(1,59)	54	403,9	57,0	2,6
Lettland	16,65	(1,26)	33	501,6	74,0	4,2
Litauen	23,09	(1,32)	11	485,4	75,2	3,3
Malaysia	22,18	(2,06)	16	520,4	74,0	2,7
Marokko	6,84	(2,02)	51	344,1	81,2	2,0
Mazedonien	24,05	(2,35)	7	467,9	87,1	2,5
Moldawien	15,80	(2,01)	35	478,5	82,4	2,5
Neuseeland	22,80	(1,40)	13	514,9	85,4	3,8
Niederlande	17,03	(2,08)	30	552,2	73,8	3,4
Norwegen	20,00	(1,68)	23	512,3	78,3	4,0
Österreich	20,80	(1,68)	21	546,9	88,7	3,3
Philippinen	16,53	(2,09)	34	376,4	93,7	2,1
Portugal	10,40	(1,05)	47	462,3	61,3	3,0
Rumänien	19,68	(2,09)	24	483,4	88,8	2,9
Russische Föderation	20,27	(1,78)	22	539,0	84,7	3,6
Schottland	26,95	(1,70)	3	504,7	86,5	3,3
Schweden	21,24	(1,59)	20	524,5	80,5	3,9
Schweiz	16,77	(1,62)	32	531,4	81,9	3,4
Singapur	18,54	(1,82)	27	613,4	83,0	2,9
Slowakische Republik	24,01	(1,26)	8	546,6	77,9	3,4
Slowenien	22,25	(1,51)	15	547,8	78,5	3,3
Spanien	15,26	(1,06)	36	498,8	68,1	3,5
Südafrika	22,05	(2,67)	17	480,6	97,3	2,2
Taiwan	27,91	(1,31)	2	592,8	90,2	2,9
Thailand	12,37	(1,60)	41	505,9	74,2	2,5
Tschechische Republik	22,45	(1,67)	14	555,7	79,5	3,9
Tunesien	6,32	(0,72)	52	452,5	56,9	2,5
Türkei	11,77	(1,35)	42	446,8	73,3	2,4
Ungarn	25,84	(1,29)	4	550,6	82,0	3,8
USA	23,13	(0,81)	10	518,9	89,2	3,5
Zypern	14,21	(0,89)	38	474,3	79,3	3,3

EFH: Geschätzter Koeffizient der Variable „zu Hause verfügbare Bücher“ (siehe Text für Modellspezifikation). Abhängige Variable: internationaler Testpunktzahl in TIMSS/TIMSS-Repeat (Mittelwert von Mathematik und Naturwissenschaften). Regressionen gewichtet mit der Erhebungswahrscheinlichkeiten der Schüler. ^a Mittelwert der fünf Kategorien: 1 (0-10 Bücher), 2 (11-25), 3 (26-100), 4 (101-200), 5 (mehr als 200 Bücher).

Quelle: Schütz et al. (2005).

Tabelle 3: Bildungspolitik und Chancengleichheit

	Spezifikation auf Länderebene		Interaktionsspezifikation auf Schülerebene mit fixen Ländereffekten		
	(1)	(2)	(3)	(4)	(5)
Zu Hause verfügbare Bücher			29,37*** (5,32)	10,61 (8,53)	30,61** (11,97)
Alter bei erster schulischer Selektion	-1,23*** (0,34)	-0,93*** (0,34)	-1,07*** (0,28)	-0,89*** (0,28)	-0,97*** (0,34)
Besuchsquote im Vorschulbereich	0,37*** (0,09)	0,21* (0,12)	0,34*** (0,09)	0,26*** (0,08)	0,04 (0,10)
Besuchsquote im Vorschulbereich (quadriert)	-0,003*** (0,001)	-0,002* (0,001)	-0,003*** (0,001)	-0,002*** (0,001)	-0,000 (0,001)
Dauer des Vorschulprogramms	-1,38* (0,78)	-1,32 (0,87)	-1,30* (0,72)	-1,55** (0,69)	-2,10** (1,02)
Bildungsausgaben pro Schüler /1000		-0,14 (0,41)		-0,14 (0,40)	-0,18 (0,40)
Bruttonationaleinkommen pro Kopf /1000		-0,16 (0,22)		-0,14 (0,23)	0,16 (0,22)
Durchschnittliche Testpunktzahl des Landes /100		4,92*** (1,71)		4,34** (1,75)	1,02 (1,65)
Anteil privater Bildungsausgaben					0,16** (0,06)
Anteil von Schülern an Privatschulen					-0,10*** (0,03)
Beobachtungen (Länder/Schüler)	53	49	295.026	276.577	156.412
Primäre Erhebungseinheiten (Länder)			53	50	27
R^2	0,400	0,268	0,414	0,417	0,294

Spezifikation auf Länderebene: entsprechend Gleichung (3). Abhängige Variable: geschätzter Effekt des familiären Hintergrundes (EFH) aus Tabelle 2. Regressionen verwenden die gewichtete Schätzmethode nach Anderson (1993) zur Berücksichtigung der geschätzten abhängigen Variable. Standardfehler in Klammern. Berichtete R^2 sind korrigierte R^2 .

Interaktionsspezifikation auf Schülerebene mit fixen Ländereffekten: entsprechend Gleichung (4). Abhängige Variable: TIMSS/TIMSS-Repeat internationale Testpunktzahl (Mittelwert aus Mathematik und Naturwissenschaften). Die angegebenen Koeffizienten beziehen sich auf die Interaktionsterme zwischen den in der ersten Spalte genannten Variablen und den zu Hause verfügbaren Büchern. Regressionen sind gewichtet mit der Erhebungswahrscheinlichkeit der Schüler innerhalb jeden Landes, und jedes Land hat das gleiche Gewicht. Clusterrobuste Standardfehler (auf Landesebene) in Klammern. Signifikanzniveaus: *** 1%. – ** 5%. – * 10%.

Quelle: Schütz et al. (2005).

Tabelle 4: Frühe Mehrgliedrigkeit, Ungleichheit und Leistungsniveau

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9) ^a
Mittelstufentest:	PISA 03	PISA 00/02	TIMSS 95	TIMSS 95	TIMSS 03	TIMSS 03	TIMSS 99	TIMSS 99	Gepoolt
Grundschulstest:	PIRLS	PIRLS	TIMSS 95	TIMSS 95	TIMSS 03	TIMSS 03	TIMSS 95	TIMSS 95	Gepoolt
Fach:	Lesen	Lesen	Mathem.	Naturw.	Mathem.	Naturw.	Mathem.	Naturw.	Gepoolt
Abhängige Variable: Standardabweichung in Mittelstufe									
Frühe Mehrgliedrigkeit	0,25** (0,11)	-0,02 (0,07)	0,15* (0,08)	0,20** (0,09)	0,01 (0,06)	0,11* (0,06)	0,004 (0,06)	0,21** (0,09)	0,10*** (0,03)
Standardabweichung in Grundschule	0,60*** (0,15)	0,25** (0,10)	0,48** (0,23)	0,85*** (0,21)	-0,02 (0,15)	0,27*** (0,10)	0,12 (0,14)	0,83*** (0,14)	0,40*** (0,06)
Konstante	2,15*** (0,45)	1,84*** (0,22)	0,81** (0,36)	0,67* (0,38)	1,33*** (0,18)	1,06*** (0,12)	1,49*** (0,25)	0,79** (0,28)	1,45*** (0,13)
R^2	0,479	0,255	0,258	0,419	0,003	0,257	0,021	0,557	0,951
Abhängige Variable: Durchschnittsleistung in Mittelstufe									
Frühe Mehrgliedrigkeit	-1,06*** (0,30)	-0,95*** (0,26)	-0,04 (0,16)	0,55*** (0,21)	-0,02 (0,17)	-0,02 (0,16)	-0,32** (0,16)	0,29 (0,25)	-0,19** (0,09)
Durchschnittsleistung in Grundschule	0,69*** (0,16)	0,66*** (0,13)	0,95*** (0,07)	0,79*** (0,10)	0,94*** (0,08)	0,97*** (0,07)	0,96*** (0,07)	0,74*** (0,10)	0,89*** (0,04)
Konstante	0,53** (0,21)	0,48** (0,19)	0,01 (0,08)	-0,17 (0,11)	0,01 (0,09)	0,01 (0,09)	0,11 (0,08)	-0,10 (0,14)	0,10 (0,13)
R^2	0,582	0,635	0,900	0,776	0,858	0,861	0,915	0,743	0,714
Länder	18	20	26	26	25	25	18	18	176

Die beiden Gleichungen in jeder Spalte wurden gemeinsam als System mithilfe Scheinbar unverbundener Regressionen (SUR) geschätzt. Standardfehler in Klammern. Signifikanzniveaus: *** 1%. – ** 5%. – * 10%. ^a Regression kontrolliert für sieben Testpaar-Dummies.

Quelle: Hanushek und Wößmann (2005) sowie eigene Berechnungen.

Ifo Working Papers

- No. 16 Wößmann, L., Ursachen der PISA-Ergebnisse: Untersuchungen auf Basis der internationalen Mikrodaten, August 2005.
- 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.