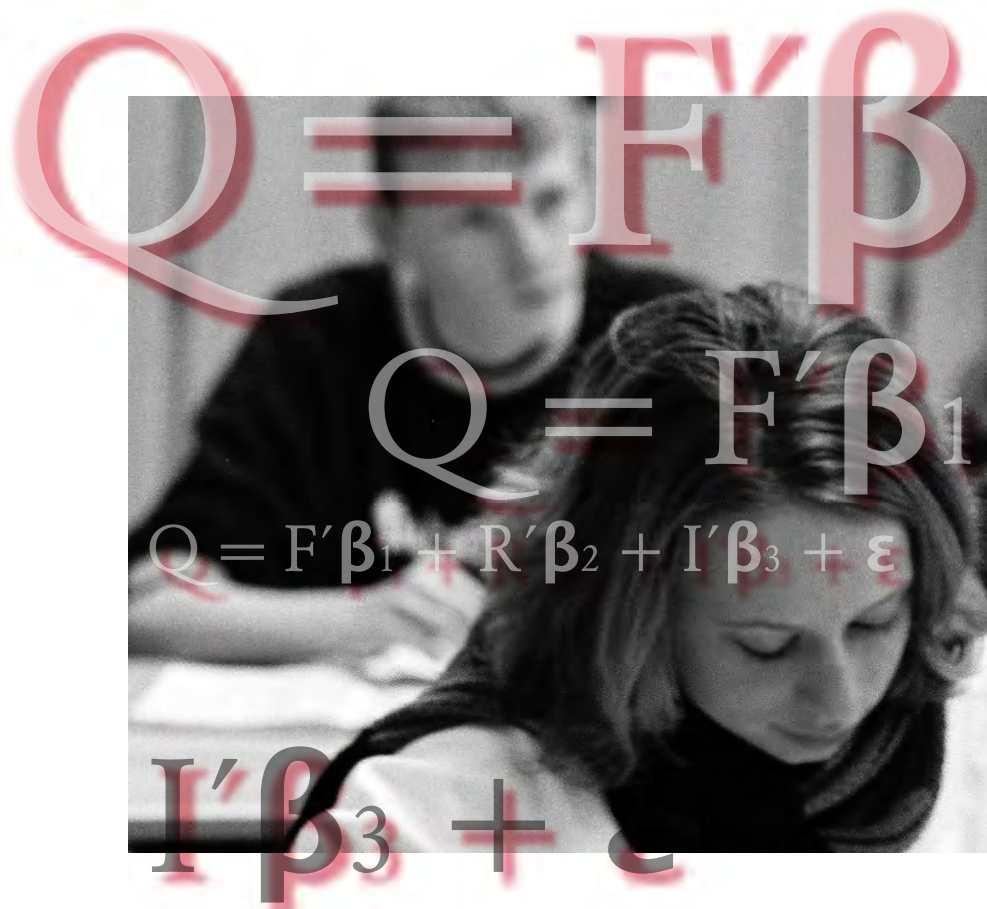


Gundel Schümer und Manfred Weiß

Bildungsökonomie und Qualität der Schulbildung

Kommentar zur bildungsökonomischen Auswertung von Daten
aus internationalen Schulleistungstudien



Impressum

Gewerkschaft Erziehung und Wissenschaft
Hauptvorstand
Reifenberger Str. 21
60489 Frankfurt
069/78973-0
Fax: 069/78973-202
E-Mail: info@gew.de
www.gew.de

Verantwortlich: Marianne Demmer, Ulf Rödde (V.I.S.d.P.)
Redaktion: Martina Schmerr
Gestaltung und Satz: Jana Roth · Kronberg
Druck: Druckerei Leutheußner, Coburg



Die Studie wurde finanziert
durch die Max-Traeger-Stiftung.

ISBN: 978-3-939470-26-7
Artikelnr.: 1302

Die Broschüre erhalten Sie im GEW-Shop (www.gew-shop.de, E-Mail: gew-shop@callagift.de,
Fax: 06103-30332-20), Mindestbestellmenge: 10 Exemplare, Einzelpreis 3,00 Euro, Preise zzgl.
Verpackungs- und Versandkosten von zurzeit 6,96 Euro brutto.

Einzelexemplare können Sie anfordern unter: broschueren@gew.de,
Fax: 069/78973-70161. Einzelpreis 4,20 Euro inkl. Versandkosten.

August 2008

Bildungsökonomie und Qualität der Schulbildung

Kommentar zur bildungsökonomischen Auswertung von Daten aus internationalen Schulleistungstudien

Inhalt	
Vorwort	5
1. Würdigung des bildungsökonomischen Forschungsbeitrags zur Qualität der Schulbildung	7
2. Der methodische Ansatz in der bildungsökonomischen Wirkungsforschung	11
3. Forschungsergebnisse zur Ressourcenwirksamkeit <i>Neuere bildungsökonomische Untersuchungen zum Einfluss der Klassengröße auf Schülerleistungen</i>	13
4. Forschungsergebnisse zur Wirksamkeit institutioneller Rahmenbedingungen	21
4.1. Zentrale Abschlussprüfungen	
4.2. Schulautonomie	
4.3. „Risiken und Nebenwirkungen“ verhaltenssteuernder Institutionen	
4.4. Wettbewerbselemente	
Leistungssteigerung durch Privatschulen?	
Sind Privatschulen leistungsfähiger?	
Schneiden Schüler in Privatschulen in PISA E besser ab?	
Neue Studien aus den USA	
5. Institutionelle Rahmenbedingungen und Chancengleichheit	37
6. Zusammenfassung	43
Literatur	47

Vorwort

In den letzten Jahren werden Schulen, Lehrkräfte, Schülerinnen und Schüler und deren Eltern des Öfteren mit Schlagzeilen konfrontiert, die auf ungläubiges Staunen bis hin zu heftigen Protesten führen. Es sind Meldungen wie zum Beispiel „Klassengrößen sind für das Lernen unerheblich“ oder „Bessere Leistungen an Privatschulen“, die sich auf die Forschungsergebnisse einflussreicher Wirtschaftsinstitute und publizitätstüchtiger Wissenschaftler stützen.

Vor allem die Behauptung, die Klassen- bzw. Gruppengröße sei ohne Einfluss auf die Qualität der Schulbildung, erregt die Gemüter, steht sie doch im deutlichen Widerspruch zu den eigenen Erfahrungen und Beobachtungen.

Allerdings ist der Einfluss von Schlagzeile und Wirtschaftswissenschaft auf die handelnden Schul- und Finanzpolitiker in der Regel stärker als die Erfahrungen der Betroffenen – insbesondere dann, wenn die Forschungsergebnisse (finanz-) politisch willkommen sind. Ökonomie schlägt Pädagogik.

Für die Gewerkschaft Erziehung und Wissenschaft (GEW) war dies Anlass, eine Expertise anzuregen, die der Frage nachgeht, wie eigentlich die Forschungsergebnisse zustande kommen, die so viel Aufsehen erregen. Wir sind froh, dass wir die Max-Traeger-Stiftung (MTS) gewinnen konnten, das Vorhaben in Auftrag zu geben. So war es möglich, mit Gundel Schümer und Manfred Weiß zwei ausgewiesene empirische Wissenschaftler zu gewinnen, die sich seit langem mit Fragen von Bildungsökonomie und Qualität der Schulbildung beschäftigen.

Neben der Darstellung des methodischen Ansatzes in der bildungsökonomischen Wirkungsforschung werden fünf Bereiche untersucht, denen Wirtschaftswissenschaftler besondere Aufmerksamkeit in ihren Analysen möglicher Einflussfaktoren der Bildungsqualität widmen: die Klassengröße, die Form von Abschlussprüfungen (zentral/dezentral), Schulautonomie und Dezentralisierung, Schulwettbewerb, Privatschulen sowie der Zusammenhang zwischen institutionellen Rahmenbedingungen und Chancengleichheit.

Die Ergebnisse der Studie sind brisant. Für alle untersuchten Bereiche kommen Schümer/Weiß (der selbst Bildungsökonom ist) zu dem Ergebnis:

„Die Bildungsökonomie hat das theoretische und methodische Arsenal der Schulqualitätsforschung erweitert. Doch basieren ihre Forschungen häufig auf wenig belastbaren Hypothesen, d.h. Hypothesen, die der Spezifik des Objektbereichs nicht angemessen Rechnung tragen. (...) Eine Auseinandersetzung mit konkurrierenden Hypothesen und konträren Forschungsergebnissen anderer Disziplinen findet in der Regel nicht statt. Dies hängt unter anderem damit zusammen, dass die meisten Wirtschaftswissenschaftler die Bildungsökonomie nicht als Teildisziplin der Bildungsforschung begreifen und in Deutschland derzeit eine institutionelle Basis für ein produktives Zusammenwirken der Bildungsökonomie mit anderen Bildungswissenschaften fehlt.... Der bildungsökonomische Produktionsfunktions-Ansatz wird der Komplexität schulischer Bildungsprozesse nicht gerecht. Der Aussagegehalt der damit gewonnenen Ergebnisse für bildungspolitisches Steuerungs Handeln, das Informationen über Wirkungsmechanismen erfordert, ist stark eingeschränkt.“

Es ist zu hoffen, dass die Studie zu intensiven fachlichen Diskussionen der beteiligten Wissenschaftsdisziplinen führt. Die Bildungsökonomie kann nur im kontinuierlichen Austausch mit den übrigen Bildungswissenschaften und der schulischen Praxis eine hilfreiche Wissenschaft für die Qualität der Schulbildung sein. Die Max-Traeger-Stiftung und die Gewerkschaft Erziehung und Wissenschaft wollen diese Diskussion befördern und unterstützen. Für handelnde Politiker möge die Studie Anregung und Hinweis sein, die Ergebnisse der Bildungsökonomie nicht ungeprüft zur Grundlage weit reichender Entscheidungen zu machen.

Marianne Demmer

Stellvertretende Vorsitzende der Gewerkschaft Erziehung und Wissenschaft

August 2008

I.

Würdigung des bildungsökonomischen Forschungsbeitrags zur Qualität der Schulbildung

**„Die Bildungsökonomie lebt, und weil sie lebt, ändert sie sich“.
Mit diesen Worten kommentierte Friedrich Edding, der Nestor der
Bildungsökonomie im deutschsprachigen Raum, Anfang der 1970er
Jahre Entwicklungen in der sich thematisch ausdifferenzierenden
Disziplin. Besser denn je charakterisieren diese Worte aktuelle
Tendenzen in der Bildungsökonomie.**

Die von einigen Kritikern (z.B. Becker 2000) schon totgesagte Disziplin erfährt gegenwärtig einen bemerkenswerten Aufschwung. Dies betrifft ihre Position innerhalb der wirtschaftswissenschaftlichen Forschung ebenso wie ihre Funktion in der Politikberatung. Unübersehbar ist die Tendenz, ihr eine wichtigere, die Politik unterstützende Rolle zuzuweisen (z.B. Sachverständigenrat 2004). Die immer knapper werdenden öffentlichen Mittel bieten dafür ebenso eine Erklärung wie der Bedeutungszuwachs, den Bildung für die Bewältigung wirtschaftlicher und gesellschaftlicher Probleme erfahren hat. Ohne Zweifel haben die bildungsökonomisch fundierten Aktivitäten der OECD maßgeblichen Anteil an dem Popularitätsschub der Disziplin. Aber auch Entwicklungen innerhalb der Bildungsökonomie haben dazu beigetragen. Zu nennen sind hier insbesondere die thematische Akzentverschiebung von theoretischen Erkenntnisinteressen der Wirtschaftswissenschaften zu Steuerungsproblemen des politisch-administrativen Systems (Heyneman 1995; Weiß 2002) sowie die verbesserte institutionelle Verankerung im Wissenschaftssystem. An fast allen außeruniversitären Wirtschaftsforschungseinrichtungen in Deutschland ist die Bildungsökonomie mittlerweile präsent. Sichtbares Zeichen der wieder erstarkten Disziplin ist die deutlich gestiegene Zahl der Mitglieder im „Bildungsökonomischen Ausschuss“ des „Vereins für Socialpolitik“. Dass die Bildungs-

ökonomie in Deutschland inzwischen international anschlussfähig geworden ist, wird eindrucksvoll durch die Publikationen von Nachwuchswissenschaftlern in hochrangigen englischsprachigen Zeitschriften und ihre Kooperation mit prominenten amerikanischen Bildungsökonomien belegt. Auch am Aufbau internationaler Netzwerke in der Bildungsökonomie sind deutsche Wirtschaftswissenschaftler maßgeblich beteiligt. Das von der Europäischen Kommission initiierte und finanzierte *European Expert Network on Economics of Education* wird vom Ifo-Institut in München koordiniert (www.education-economics.org). Diesem Netzwerk gehören inzwischen fast 350 Wissenschaftler an.

Im Zuge ihrer programmatischen Neuausrichtung hat die Bildungsökonomie in letzter Zeit in verstärktem Maße Themen besetzt, die traditionell in den Zuständigkeitsbereich anderer Bildungswissenschaften fallen. Besonderes Augenmerk gilt dabei Fragen der Qualität der Schulbildung. Auslöser dafür ist nicht nur die hohe öffentliche Aufmerksamkeit, die diese Thematik im Gefolge von TIMSS und insbesondere PISA erfahren hat, sondern auch spektakuläre Befunde neuerer empirischer Studien zum Wirtschaftswachstum. Sie identifizieren auf Schülerleistungen basierende Merkmale der Humankapitalqualität als besonders wichtigen Wachstumstreiber (z.B. Hanushek & Kimko

2000; Hanushek 2005; Hanushek & Wößmann 2007). Dies hat das Forschungsinteresse der Bildungsökonomie verstärkt auf aussichtsreiche strategische Ansatzpunkte einer gezielten bildungspolitischen Beeinflussung der Humankapitalqualität im Schulbereich gelenkt. Die aus den großen internationalen Schulleistungsvergleichen vorliegenden Datensätze bieten für darauf gerichtete Analysen besonders günstige Voraussetzungen. Das von der Bildungsökonomie in diesem Kontext bearbeitete Themenspektrum reicht von Fragen der Ressourcenauswirkung und der Effekte von Dezentralisierung, Autonomie und Wettbewerb bis hin zu originären pädagogischen Fragen wie



Die Anwendung anspruchsvoller Analysemethoden ist gewissermaßen das Markenzeichen der „neuen“ Bildungsökonomie.

etwa der Beeinflussung von Schulleistungen durch das Einschulungsalter, Klassenwiederholungen, die Zusammensetzung der Schülerschaft, leitungsdifferenzierten Unterricht und zentrale Prüfungen. Mit Untersuchungen zum moderierenden Einfluss schulstruktureller Faktoren auf den Kompetenzerwerb von Schülern¹ und seine Abhängigkeit von ihrem sozioökonomischen und soziokulturellen Hintergrund wendet sich die bildungsökonomische Forschung neuerdings auch Fragen der Chancengleichheit im Bildungswesen zu.

Die Bildungsökonomie dringt damit in Bereiche vor, die andere Disziplinen als ihr angestammtes Forschungsterrain betrachten und für die sie einen Expertisevorsprung reklamieren. Schnell wird darauf von den betroffenen Disziplinen mit dem Vorwurf unproduktiver Doppelarbeit oder fehlender Kompetenz der Bildungsökonomie auf Grund unzureichender Institutionenkenntnisse reagiert. Zumindest der erste Punkt lässt sich mit dem Hinweis entkräften, dass die in einem Theorie- und Methodenwettbewerb stehende Arbeit verschiedener Disziplinen an einem gemeinsamen Forschungsgegenstand eine wichtige Voraussetzung für Erkenntnisfortschritt ist. Mit dem Rekurs etwa auf Erklärungsansätze aus der *Neuen Institutionenökonomie*² werden von der Bildungsökonomie neue theoretische Zugangswege bei der Erforschung der Leistungswirksamkeit differenter Steuerungssysteme im Bildungsbereich erschlossen. Zugleich hat sie das Methodenarsenal in der Schulqualitätsforschung um neue Verfahren aus der Ökonometrie erweitert. Sie kann z.B. zeigen, dass es Alternativen zu Modellen gibt, die in der erziehungswissenschaftlichen Forschung üblicherweise zur Analyse hierarchischer Daten eingesetzt werden (HLM-Modelle), und dass sich unter bestimmten Voraussetzungen Daten aus Stichprobenerhebungen quasi-experimentell auswerten lassen (vgl. Schneider, Carnoy u.a. 2007).

Die Anwendung anspruchsvoller Analysemethoden ist gewissermaßen das Markenzeichen der „neuen“ Bildungsökonomie. Freilich ist

1. Im Interesse der besseren Lesbarkeit des Texts stehen die Bezeichnungen „Schüler“ und „Lehrer“ jeweils für beide Geschlechter.

2. Die Neue Institutionenökonomie beschäftigt sich mit den Auswirkungen von Institutionen (z.B. Verfügungsrechten, Verträgen, Organisationsstrukturen, Märkten) auf menschliches Verhalten. Sie untersucht insbesondere Möglichkeiten der effizienten Gestaltung von Institutionen. Die der Neuen Institutionenökonomie zurechenbaren Ansätze (Theorie der Verfügungsrechte, Transaktionskostentheorie und Principal Agent Theory) sind durch weitgehend übereinstimmende Annahmen zum menschlichen Verhalten gekennzeichnet: individuelle Nutzenmaximierung und begrenzte Rationalität des Handelns (vgl. Picot, Dietl & Franck 1999, S. 54 ff.).

dies keine hinreichende Bedingung für Erkenntnisfortschritt. Dazu bedarf es – neben einer tragfähigen Datenbasis – auch begründeter Hypothesen, die in angemessener Weise der Spezifik des Forschungsgegenstandes Rechnung tragen. Wie beispielhaft zu zeigen sein wird, liegen hier die wesentlichen Defizite der bildungsökonomischen Forschung zur Qualität der Schulbildung. Eine kritische Würdigung ihres Beitrags zur Verbesserung des Forschungsstandes in diesem Bereich und zur Fundierung bildungspolitischen Steuerungshandelns wird auch nicht darüber hinwegsehen können, dass relevante Befunde anderer Bildungswissenschaften von der Bildungsökonomie bislang entweder gänzlich ignoriert oder höchst selektiv – zur Absicherung der eigenen Hypothesen und Befunde – zur Kenntnis genommen werden. Eine Auseinandersetzung mit konkurrierenden Hypothesen und konträren Forschungsergebnissen anderer Disziplinen findet in der Regel nicht statt. Karrierezwänge, die das Publikationsverhalten beeinflussen, bieten

dafür sicherlich ebenso eine Erklärung wie die geringe Neigung vieler Wirtschaftswissenschaftler, die Bildungsökonomie als Teildisziplin der Bildungsforschung zu begreifen und sich auf eine Mitarbeit in einem institutionalisierten multidisziplinären Forschungsverbund einzulassen. Bildungsökonomische Forschung wird in Deutschland mittlerweile fast nur noch in wirtschaftswissenschaftlichen Fachbereichen und in Einrichtungen der außeruniversitären Wirtschaftsforschung betrieben. In Bildungsforschungsinstituten wie dem Max-Planck-Institut für Bildungsforschung und dem Deutschen Institut für Internationale Pädagogische Forschung, in denen die Bildungsökonomie in der Vergangenheit eine bedeutende Rolle spielte, ist sie nicht einmal mehr rudimentär vertreten. Für eine Synergieeffekte sichernde interdisziplinäre Kommunikation und Kooperation fehlt derzeit eine institutionelle Basis. In den Empfehlungen des Wissenschaftsrats zum Ausbau der Bildungsforschung (Wissenschaftsrat 2001, S. 54 ff.) war dies einmal anders vorgesehen.

2.

Der methodische Ansatz in der bildungsökonomischen Wirkungsforschung

Das primäre Interesse der bildungsökonomischen Wirkungsforschung gilt bildungspolitischen Gestaltungsparametern. Bei der empirischen Untersuchung ihres Einflusses auf den Schulerfolg bedient sich die Bildungsökonomie vorrangig des Konzepts der Bildungsproduktionsfunktion; die untersuchten Variablenzusammenhänge werden darin als Input-Output-Beziehungen modelliert.

Drei Gruppen von Einflussfaktoren stehen im Mittelpunkt der Betrachtung: Merkmale des familiären Hintergrunds der Schüler, Merkmale der Ressourcen der Schulen und institutionelle Merkmale der Bildungssysteme wie etwa der Grad der Zentralisierung oder Dezentralisierung in wichtigen Handlungsfeldern (Curricula, Personal, Finanzen, Definition von Leistungsstandards) und die paradigmatische Orientierung des Steuerungssystems (Wettbewerb vs. staatlich-bürokratische Steuerung).

Das üblicherweise der Effektschätzung zugrunde liegende Modell einer Bildungsproduktionsfunktion lässt sich formal wie folgt darstellen:

$$Q = F'\beta_1 + R'\beta_2 + I'\beta_3 + \epsilon,$$

mit Q als Qualitätsindikator (meist Testleistungen der Schüler), F' als Vektor von Merkmalen des familiären Hintergrunds, R' als Vektor von Merkmalen der schulischen Ressourcen, I' als Vektor von institutionellen Merkmalen der Bildungssysteme und ϵ als Fehlerterm. Die Parametervektoren β_1 bis β_3 , die die Stärke des Zusammenhangs zwischen dem Qualitätsindikator (Testleistungen) und den jeweiligen Einflussfaktoren (unabhängigen Variablen) widerspiegeln, werden in einer multivariaten Regressionsanalyse geschätzt. Die Effektschätzung kann dabei für

verschiedene Analyseeinheiten vorgenommen werden: Schüler, Klassen, Schulen, Schulbezirke, Länder. „Multivariat“ bedeutet, dass bei der Schätzung des Zusammenhangs der mittleren Schülerleistungen mit jeder einzelnen Variablen der Einfluss aller übrigen Variablen konstant gehalten wird; z.B. wird der Zusammenhang mit der Variablen „Schulautonomie“ so berechnet, als hätten die Schüler aller Schulen den gleichen familiären Hintergrund.

„Damit wird eine Konfundierung der jeweiligen Effekte mit anderen beobachteten Hintergrundeffekten ausgeschlossen“ (Wößmann 2005c, S. 153).

Das Modell der Bildungsproduktionsfunktion ist dadurch gekennzeichnet, dass ein *direkter Wirkungszusammenhang* zwischen den Einflussfaktoren (unabhängigen Variablen) und der abhängigen Variablen (Testleistungen) unterstellt wird: dass Ressourcen und institutionelle Kontextbedingungen Lernergebnisse unmittelbar beeinflussen. Die Annahme eines direkten Einflusses distaler, vom Unterrichtsgeschehen entfernter Faktoren ist theoretisch nicht zu begründen, sondern nur ein über unterrichtsnahe prozessuale Bedingungsfaktoren vermittelter Einfluss auf Schulleistungen: das Angebot an Lerngelegenheiten, die Qualität der Instruktion und die Nutzung der Lerngelegenheiten durch die Schüler (z.B. Kilgore & Pendleton 1993;

Creemers & Reezigt 1996, S. 208 ff.; Fend 1998, S. 268 ff.). In dem Produktionsfunktions-Ansatz der Bildungsökonomie bleibt dies als *black box* ausgeblendet. Wird z.B. ein Zusammenhang zwischen bestimmten institutionellen Merkmalen und Schülerleistungen festgestellt, dann wird dies als Ergebnis des

verhaltenssteuernden Einflusses von Anreizstrukturen gesehen; das Verhalten der schulischen Akteure selber interessiert nicht. Auf die Problematik dieser Restriktion wird später noch ausführlich eingegangen (vgl. Abschnitt 4.3).

3.

Forschungsergebnisse zur Ressourcenwirksamkeit

Unter Anwendung des Produktionsfunktionsansatzes ist in der Vergangenheit von der Bildungsökonomie vor allem die Frage empirisch untersucht worden, welchen Einfluss vorfindbare Unterschiede in der Ressourcenausstattung der Schulen auf Schülerleistungen haben, und zwar unter Konstanthaltung des Einflusses aller anderen berücksichtigten Variablen.

Die Ergebnisse dieser Untersuchungen sind in verschiedenen Forschungsübersichten zusammengefasst worden. Besondere Publizität haben die mehrfach aktualisierten Auswertungen US-amerikanischer Studien durch Eric Hanushek erlangt. Sein Resümee repräsentiert die derzeit herrschende Meinung in der Bildungsökonomie: „With over three decades of analysis, new studies have reinforced earlier conclusions: Today’s schools exhibit continuing inefficiency in their operations as there is no strong or consistent relationship between variations in school resources and student performance” (Hanushek 1997, S. 141).

Das Fazit aus den Befunden US-amerikanischer Produktivitätsstudien korrespondiert weitgehend mit dem einer separaten Länderauswertung von TIMSS-Daten durch Hanushek und Luque (2003). Real vorfindbaren Unterschieden in der Ausstattung mit finanziellen, personellen und materiellen Ressourcen kommt – zumindest in Industriestaaten – offenbar nur geringe Bedeutung als Varianzquelle von Schülerleistungen zu. Aus diesem Ergebnis kann nun nicht auf eine grundsätzliche Bedeutungslosigkeit schulischer Ressourcen für die Qualität der Schulbildung geschlossen werden, denn: Erstens kann von einer Effektabhängigkeit vom Ressourcenniveau ausgegangen werden (Schwellenwert-Effekte, abnehmende Grenzerträge mit zunehmendem

Ressourceneinsatz); zweitens decken die Studien nicht das gesamte Spektrum an schulischen Ressourcen und Möglichkeiten der Mittelverwendung ab (vgl. zu weiteren methodischen Kritikpunkten von Produktivitätsstudien Weiß & Timmermann 2004).

Die – kontrovers geführte – Debatte um die Wirksamkeit schulischer Ressourcen hat sich vor allem an divergierenden Vorstellungen von Wissenschaft und Praxis über die Wichtigkeit der *Klassengröße* als Determinante der Unterrichtsqualität entzündet. Dieses *Evergreen* bildet auch den Gegenstand neuerer bildungsökonomischer Analysen der Datensätze internationaler Schulleistungstudien.

Neuere bildungsökonomische Untersuchungen zum Einfluss der Klassengröße auf Schülerleistungen

Seit vielen Jahren beschäftigt man sich in der Pädagogik, der Psychologie und der Bildungsökonomie mit der Frage, wie sich Klassenfrequenzen auf Schüler auswirken. Die ersten empirischen Untersuchungen des Problems wurden bereits Anfang des vergangenen Jahrhunderts durchgeführt. Seither sind nicht nur viele tausend wissenschaftliche Arbeiten erschienen, in denen die Effekte unterschiedlich großer Lerngruppen gemessen werden, sondern



auch eine kaum mehr überschaubare Menge von Literaturberichten und Meta-Analysen, mit deren Hilfe versucht wird, eine gewisse

Eine Kontrolle der Klassenzusammensetzung ist notwendig, da die Klassenfrequenzen in den meisten Schulsystemen nicht unabhängig von den Fähigkeiten, den Leistungen und dem Verhalten der Schüler sind.

Ordnung in die Ergebnisse der vorliegenden Studien zu bringen. Schließlich sind in den letzten Jahren einige komparative Untersuchungen dazugekommen, in denen Klassengrößeneffekte in verschiedenen Ländern auf der Basis von Daten aus den jüngsten internationalen Schulleistungsstudien analysiert werden. Zu ihnen gehören bildungsökonomische Arbeiten von Martin West und Ludger Wößmann (West & Wößmann 2006, Wößmann 2005b, Wößmann & West 2006; siehe auch Wößmann 2005a und 2005c)³, die – wie viele andere Beiträge zum Thema „Klassengröße“ – kontroverse Diskussionen ausgelöst haben. All diejenigen, die an die hohen Kosten niedriger Klassenfrequenzen denken müssen, werden durch die Ergebnisse der Untersuchungen beruhigt, denn diese bestätigen die Befunde früherer Studien, nach denen das Lernen in kleinen Klassen in der Regel allenfalls geringfügige positive Auswirkungen auf die Schülerleistungen hat. Die Befürworter kleiner Klassen, die häufig zunächst einmal an die Arbeits- und Lernbedingungen von Lehrern und Schülern denken, verweisen dagegen auf andere Studien,

in denen mehr oder weniger große positive Effekte kleiner Klassen oder Lerngruppen nachgewiesen wurden (z.B. das Projekt STAR und die Londoner *Class Size Study*; vgl. zusammenfassend Brahm 2006). Folglich ist zu fragen, welche Bedeutung den bildungsökonomischen Arbeiten zukommt, die sich auf die Daten der internationalen Schulleistungsvergleiche stützen.

Die Analysen von West und Wößmann beziehen sich auf Daten aus den IEA-Studien TIMSS und TIMSS-Repeat (Martin & Kelly 1997). Anders als in den PISA-Studien, die Stichproben 15-jähriger Schüler untersuchen, wurden in TIMSS ganze Klassen getestet, und zwar nahm in den zur Stichprobe gehörenden Schulen jeweils eine Klasse der beiden Klassenstufen mit den größten Anteilen 13-jähriger Schüler an der Erhebung teil; in der Regel handelte es sich um eine siebte und eine achte Klasse. Ferner waren die in TIMSS verwendeten Tests der mathematischen und naturwissenschaftlichen Leistungen auf die Lehrpläne der beteiligten Länder zugeschnitten. Schließlich war auch ein Lehrerfragebogen eingesetzt worden, der Informationen zu den Lehrern der untersuchten Fächer und Klassen bereitstellte. Aufgrund dieser Merkmale des Untersuchungsdesigns schien es den Autoren möglich, Effekte der Klassengröße von Effekten der Klassenzusammensetzung zu trennen und die Testleistungen der Schüler als Ergebnisse des Unterrichts zu interpretieren, den sie im Schuljahr der Datenerhebung von ihrem Lehrer in ihrer Klasse erhalten haben.

Eine Kontrolle der Klassenzusammensetzung ist notwendig, da die Klassenfrequenzen in den

3. Die meisten dieser Studien sind mehrmals als Arbeitspapiere erschienen, ehe sie in Zeitschriften veröffentlicht wurden. Die vorliegende Arbeit bezieht sich ausschließlich auf die Zeitschriftenbeiträge und zitiert, soweit möglich, aus den deutsch verfassten Texten.

meisten Schulsystemen nicht unabhängig von den Fähigkeiten, den Leistungen und dem Verhalten der Schüler sind. In der Regel sind Klassen in Schulen, die von schwachen Schülern besucht werden, etwas kleiner als in Schulen mit leistungsstarken Schülern, da man versucht, die Benachteiligung schwacher Schüler, die sich in Schulen bestimmter Schulformen oder bestimmter Schuleinzugsbereiche konzentrieren, durch etwas günstigere Lernbedingungen zu kompensieren. Entsprechendes gilt für die schulinterne Bildung von Lerngruppen mit unterschiedlichem Leistungsniveau. Aufgrund solcher kompensatorischen Bemühungen sind Klassengrößen in Bezug auf die Schülerleistungen häufig endogen. Das bedeutet, dass die Qualität von Studien zur Auswirkung unterschiedlicher Klassenfrequenzen auf die Schülerleistungen ganz entscheidend davon abhängt, ob die potenzielle Endogenität der Klassengrößen geprüft bzw. bei der Berechnung der Klassengrößeneffekte kontrolliert wird.

Um sicherzustellen, dass die Klassenfrequenzen nicht mit der kognitiven oder sozialen Zusammensetzung der Schülerschaft der einzelnen Schulen oder Klassen zusammenhängen, wählen West und Wößmann für ihre Analysen ein zweistufiges Verfahren (siehe die ausführliche Darstellung in Wößmann 2005b.) Zunächst berücksichtigen sie so genannte fixe Schuleffekte, d.h. Merkmale der Schulen, die zur Bildung verschieden großer Klassen führen, wie beispielsweise hohe Anteile von Migranten oder von Schülern aus bildungsfernen Schichten. Anschließend haben sie es nur noch mit schulinternen Schwankungen der Klassenfrequenzen zu tun, die allerdings ebenfalls endogen oder leistungsabhängig sein können. Um exogene Unterschiede in den Klassengrößen zu

identifizieren, machen sich West und Wößmann dann die beiden folgenden Sachverhalte zunutze:

1. Natürliche Kohortenschwankungen:

Da die von den einzelnen Schulen aufgenommenen Schülerjahrgänge nicht immer gleich stark sind, kommt es auf verschiedenen Klassenstufen zu mehr oder weniger großen Schwankungen der Klassenfrequenzen, die nichts mit den Leistungen der Schüler zu tun haben und sich daher – dem Vorgehen von Hoxby (2000) entsprechend – zur Untersuchung von Effekten der Klassenfrequenzen auf die Schülerleistungen nutzen lassen.

2. Regelinduzierte Diskontinuitäten:

Im öffentlichen Schulwesen der meisten Länder ist administrativ geregelt, wie groß die Klassen in den verschiedenen Schulformen bzw. Schuleinzugsbereichen maximal sein dürfen und ab welcher Schülerzahl Klassen geteilt werden müssen. Dementsprechend kann es schulintern zu großen exogenen Unterschieden in den Klassenfrequenzen verschiedener Klassenstufen kommen. Beispielsweise kann ein Schülerjahrgang aus einer Klasse mit 30 Schülern bestehen, ein anderer aus zwei Klassen mit 15 bzw. 16 Schülern. Wie Angrist und Lavy (1999) gezeigt haben, sind derlei regelabhängige Differenzen der Klassengrößen gut dazu geeignet, Effekte der Klassenfrequenzen auf die Schülerleistungen zu untersuchen.

Da in TIMSS Daten zur durchschnittlichen Größe der siebten und achten Klassen der einzelnen Schulen erhoben wurden und da für jedes Land die durchschnittlichen Leistungen der Schüler jeweils beider Klassenstufen berechnet wurden, verfügen West und Wößmann über so genannte Instrumentvariablen⁴, die die Feststellung der Effekte exogener Unterschiede in den Klassenfrequenzen ermöglichen: Die vom durchschnittlichen Leistungszuwachs von der siebten zur achten Klassenstufe abweichende Differenz der Leistungen der Siebt- und Achtklässler einer Schule lässt sich auf Unterschiede in ihren Klassenfrequenzen zurückführen. Die regelinduzierten Diskontinuitäten der Klassenfrequenzen bieten außerdem die Möglichkeit, die Leistungen von Schülern in kleinen und großen Klassen miteinander zu vergleichen.

Mithilfe der geschilderten Verfahren konnten West und Wößmann belegen, dass niedrige Klassenfrequenzen in den meisten Ländern bzw. Schulsystemen zur Kompensation von Leistungsschwächen genutzt werden (West & Wößmann 2006). Sie untersuchten dann, ob die verbleibenden, bezüglich der Schülerleistungen exogenen Schwankungen der Klassenfrequenzen mit Unterschieden in den Testleistungen korrespondieren und wie stark diese mit der Größe der untersuchten Klassen variieren. Das Niveau ihres methodischen Vorgehens ist beeindruckend. Ihre Analysen beziehen sich freilich nur auf 11 bzw. 18 verschiedene Schulsysteme (Wößmann & West 2006 bzw. West & Wößmann 2006 und Wößmann 2005b), da sinnvollerweise lediglich Systeme betrachtet wurden, aus denen für mindestens 50 Schulen die relevanten Daten für beide Klassenstufen

vorlagen. Dass die Ergebnisse komplexer statistischer Berechnungen bei diesen Fallzahlen nicht so leicht zufallskritisch abzusichern sind, liegt auf der Hand.

Schon aus diesem Grund ist es nicht verwunderlich, dass die Ergebnisse ihrer Analysen „keinerlei Evidenz“ dafür liefern, dass kleinere Klassen zu höheren Leistungen führen (Wößmann 2005c, S. 14). Nur in jeweils zwei der elf von ihnen untersuchten Länder, die an TIMSS teilnahmen, hatten die Schüler kleiner Klassen signifikant bessere Leistungen in Mathematik bzw. den Naturwissenschaften. Bei der Zusammenfassung der Testleistungen beider Bereiche waren die Schüler kleiner Klassen wiederum nur in zwei Ländern – Griechenland und Island – signifikant besser als die Schüler großer Klassen (Wößmann & West 2006). Die erneute Untersuchung der elf Länder zusammen mit sieben, die an TIMSS-Repeat teilgenommen hatten, führte im Wesentlichen zum gleichen Ergebnis: Auch hier hatten kleine Klassen nur in Griechenland und Island günstige Wirkungen auf die Schülerleistungen, während in elf Ländern positive Effekte auszuschließen waren (West & Wößmann 2006). Dass die Leistungen der griechischen und isländischen Schüler in TIMSS unter dem internationalen Durchschnitt lagen, während die Schüler in Singapur, dem Land mit den größten Klassen, am besten abschnitten, ist für die Autoren ein Hinweis darauf, dass kleine Klassen für die Effektivität von Schulsystemen nicht erforderlich sind. Überlegungen zur Bedeutung des kulturellen Kontexts der verschiedenen Schulsysteme werden nicht angestellt.

4. Instrumentvariablen korrelieren hoch mit den Erklärungsvariablen, haben jedoch einen vernachlässigbaren Einfluss auf die abhängigen Variablen. Im vorliegenden Fall korrelieren die durchschnittlichen Klassengrößen der beiden Schülerjahrgänge hoch mit den Größen der untersuchten Klassen, sind im Hinblick auf die Leistungen der Schüler aber exogen.

Vielleicht waren solche Überlegungen aber der Grund für Wößmann, eine weitere Studie durchzuführen, die sich auf die Analyse kulturell besser vergleichbarer Länder beschränkt, und zwar auf 17 europäische Länder (ohne ehemalige sozialistische Republiken) und die USA (Wößmann 2005b)⁵. Außerdem hat Wößmann die deutsche TIMSS-Stichprobe im Hinblick auf die Leistungsrelevanz unterschiedlicher Klassenfrequenzen untersucht (Wößmann 2005c). Die vergleichende Studie ergab, dass sich kleine Klassen nur in dreien der untersuchten 18 Länder günstig auf die Schülerleistungen auswirkten; die Effekte waren signifikant, aber keineswegs stark. Schwache positive Wirkungen, die aber nicht signifikant waren, zeigten sich in 14 Ländern. In einem weiteren Land, nämlich der Schweiz, waren – trotz der Kontrolle der Endogenität der Klassenfrequenzen – die Schülerleistungen in großen Klassen signifikant besser. Die Analyse der deutschen Daten ergab ebenfalls, dass sich „der kausale Effekt der Klassengröße auf die Schülerleistungen [...] statistisch nicht signifikant von null unterscheiden lässt“, d.h. dass „Verkleinerungen der Klassen“ in Deutschland „kaum zu nennenswerten Verbesserungen der mathematischen Leistungen der Schüler“ führen würden (S. 13). Wößmann zieht daraus den Schluss, dass niedrigere Klassenfrequenzen – aufgrund ihrer bescheidenen Effekte – in Deutschland ebenso wenig wie in anderen westeuropäischen Ländern ökonomisch zu rechtfertigen wären, da sie wesentlich höhere Kosten als die bisherigen Klassengrößen verursachen würden (S. 14). So weit so gut, könnte man sagen, wenn Wößmann sich darauf beschränkt hätte, Kosten und Nutzen kleiner Klassen im Hinblick auf die

Leistungen von Sekundarschülern in Mathematik und den Naturwissenschaften gegeneinander abzuwägen, anstatt die Ergebnisse seiner Analysen zu verallgemeinern und Vorschläge zur Reform des Schulwesens zu machen. Darauf wird noch zurückzukommen sein.

Zunächst ist zu fragen, ob das methodische Vorgehen von West und Wößmann einer kritischen Prüfung standhält. Wie bereits angemerkt, beeindrucken die Studien durch das Niveau ihrer statistischen Datenverarbeitung. Dennoch ist es nicht möglich, die Validität der Ergebnisse einzuschätzen:

- Der Hinweis, dass „ein kompletter Satz fixer Schuleffekte“ berücksichtigt wird (Wößmann 2005c, S. 11), um bei der Berechnung von Effekten der Klassenfrequenzen auf die Leistungen Verzerrungen aufgrund endogener Klassengrößen auszuschließen, hilft dem Leser nicht viel, wenn er die Variablen nicht kennt.
- Effektschätzungen, die fixe Effekte berücksichtigen, führen im Allgemeinen zu Informationsverlusten und damit auch zu Unterschätzungen der tatsächlich vorhandenen Effekte (siehe Schneider, Carnoy u.a. 2007).
- Die Berücksichtigung fixer Schuleffekte schützt nicht vor Fehlinterpretationen, da die meisten Schüler auf ihren zuvor besuchten Grundschulen unter ganz anderen Bedingungen gelernt haben und da mit Effekten von regionalen oder kulturspezifischen Einflussgrößen zu rechnen ist, die in TIMSS ebenfalls nicht erfasst wurden.
- Natürliche Kohortenschwankungen sind in der Regel nicht besonders groß. Bei der ver-

5. Die Zahl der TIMSS-Länder, aus denen für mindestens 50 Schulen die relevanten Daten für beide Klassenstufen vorlagen, konnte aufgrund der Imputation fehlender Werte erhöht werden. Dies machte die Beschränkung auf die westeuropäischen Teilnehmerstaaten möglich. Der Nachteil von Analysen mit imputierten Durchschnittswerten ist, dass tatsächlich vorhandene Effekte unterschätzt werden.

gleichsweise geringen Zahl von Schulen – und das heißt auch von siebten bzw. achten Klassen pro Land – ist nicht mit signifikanten und schon gar nicht mit starken Klassengrößeneffekten zu rechnen. In sieben der 18 Länder waren präzise Schätzungen mithilfe dieses Verfahrens überhaupt nicht möglich (siehe Wößmann 2005b, S. 471).

- Regelinduzierte Diskontinuitäten der Klassenfrequenzen kommen in den relativ kleinen Schulstichproben der Länder nicht besonders häufig vor. Demgemäß war in vier Ländern die Basis für verlässliche Effektschätzungen nicht vorhanden (vgl. Wößmann 2005b, S. 471).
- Die vergleichende Interpretation der Ergebnisse der verschiedenen Länder ist problematisch. Dass einzelne der zur Analyse eingesetzten Verfahren⁶ für insgesamt 14 der 18 untersuchten Länder günstige Effekte kleiner Klassen ergeben, die aber nur in drei Ländern signifikant sind, und dass etliche Verfahren in zehn Ländern keine genauen Effektschätzungen erlauben, kann als Beleg für die geringe Bedeutung von Klassenfrequenzen hinsichtlich der Schülerleistungen gewertet werden. Es könnte freilich auch ein Hinweis darauf sein, dass die verwendeten Daten und Methoden zur Beurteilung der Leistungsrelevanz unterschiedlicher Klassengrößen nicht ausreichen.

West und Wößmann analysieren TIMSS-Daten, die sich auf die mathematischen und naturwissenschaftlichen Leistungen 13-Jähriger und auf Merkmale ihrer Lernbedingungen beziehen. Ihre Arbeiten lassen aber erkennen, dass sie ihren Befunden wesentlich allgemeine Bedeutung beimessen. Zwar wird gelegentlich angesprochen, dass kleine Klassen für Schulanfänger oder Vorschulkinder wichtiger sein mögen als für 13-Jährige oder dass Grundschullehrer eventuell unter anderen Bedingungen arbeiten als Lehrer auf der Sekundarstufe I. Es wird auch geprüft, ob Klassenfrequenzen – abhängig von der sozialen Herkunft der Schüler und dem Bildungsniveau ihres Elternhauses – differenzielle Wirkungen auf die Schülerleistungen haben.⁷ Schließlich wird auch untersucht, ob die länderspezifischen Klassengrößeneffekte mit den Qualifikationen der Lehrer zusammenhängen.⁸ Aber weitere Relativierungen und Problematisierungen der Ergebnisse fehlen. Beispielsweise wird nicht reflektiert, dass die Befunde fachspezifisch sein könnten⁹. Ferner wird nicht thematisiert, nach welchen Unterrichtsprinzipien, nach welchen Methoden und mit welchen Medien die Schüler unterrichtet wurden, d.h. dass die Leistungsrelevanz kleiner Klassen von der Unterrichtsgestaltung und den Unterrichtszielen abhängig sein könnte und daher im Zusammenhang mit Unterrichtsprozessen untersucht werden müsste. Es kommt auch nirgendwo zur Sprache, dass Testleistungen zwar mit Schulleistungen zusammenhängen, aber nicht mit ihnen gleichzusetzen sind, geschweige denn, dass hohe Schulleistungen nur eines von mehreren Lern-

6. Vgl. die Zusammenfassung der Ergebnisse, bei der Wößmann bis zu fünf verschiedene Effektschätzungen pro Land darstellt (Wößmann 2005c, S. 471).

7. Das Ergebnis ist wider Erwarten negativ.

8. Die Lehrerqualifikationen werden erstaunlicherweise nicht nur bei der Untersuchung der einzelnen Länder sondern auch beim Ländervergleich mit der Höhe der Lehrergehälter gleichgesetzt.

9. Dafür gibt es inzwischen empirische Belege: Wie die DESI-Studie gezeigt hat, bieten kleine Klassen im Englischunterricht mehr Kommunikationsmöglichkeiten, die sich signifikant positiv auf das Hörverständnis von Schülern auswirken (DESI-Konsortium 2007)

zielen sind, die in der Schule prinzipiell immer zusammen mit guten Leistungen angestrebt werden müssen.

Besonders problematisch ist aber sicher, dass bei der Interpretation der Ergebnisse nicht berücksichtigt wird, dass Schülerleistungen im allgemeinen Ergebnisse *langjähriger* Lernprozesse *in* und *außerhalb* der Schule sind. Manche Schüler haben vielleicht jahrelang in kleinen Klassen gesessen, ehe sie in die in TIMSS untersuchte Klasse kamen, manche haben jahrelang viel Unterricht in Mathematik oder den Naturwissenschaften gehabt, oder sie haben jahrelang viel Hausaufgaben gemacht oder auch viel Nachhilfeunterricht bekommen. Das heißt, ihre Testleistungen werden keineswegs allein vom Unterricht im Schuljahr der Datenerhebung beeinflusst. Sofern Daten zur Lerngeschichte der Schüler erhoben wurden, was in TIMSS allerdings nicht der Fall war, und sofern sich vergleichbare Schüler in bestimmten Schulen konzentrieren, kommt es, wenn fixe Schulleffekte berücksichtigt werden, bei den nationalen Analysen der Daten zu Klassenfrequenzen und Schülerleistungen *nicht* zu schwerwiegenden Problemen. Bei der vergleichenden Interpretation der Befunde für die verschiedenen Länder ist dies anders, da es länderspezifische institutionelle Bedingungen des Lernens und damit auch länderspezifische Lernprozesse und Schulerfahrungen gibt, die den Einfluss der Klassenfrequenzen auf die Leistungen zur Zeit der Datenerhebung für TIMSS verstärken, neutralisieren oder abschwächen können. Bei den Ländervergleichen vermisst man im Übrigen auch Hinweise auf die sozialen und kulturellen Kontexte, in denen in verschiedenen Ländern gelernt und unterrichtet wird. Wie sich große oder kleine Klassen auf die Schülerleistungen auswirken, hängt schließlich auch mit kultur-

spezifischen Erziehungs- und Unterrichtsstilen und mit der kulturspezifischen Bedeutung des Lernens in und außerhalb der Institution „Schule“ zusammen. Wer all diese Randbedingungen ignoriert, braucht sich über schwache, nicht signifikante Wirkungen der Klassengröße auf die Schülerleistungen nicht zu wundern.

Aus pädagogischer Sicht werfen die Studien von West und Wößmann mehr Fragen auf als sie beantworten. Wahrscheinlich kann man Mathematikaufgaben tatsächlich in großen und kleinen Klassen mit dem gleichen Erfolg in Stillarbeit lösen lassen, aber: Sollte die Klassengröße tatsächlich keine Rolle spielen, wenn es z.B. darum geht, dass Schüler eine fremde Sprache sprechen lernen oder lernen sollen, sich frei zu einem Thema zu äußern? Könnte es sein, dass Lehrer bei den üblichen Klassenfrequenzen auf progressiven Unterricht verzichten und dass die relativ geringfügigen Differenzen in den Klassengrößen nur deshalb ohne größere Bedeutung für die Schülerleistungen sind? Könnte es sein, dass es sich nicht an den getesteten Leistungen von Schülern ablesen lässt, ob sie beispielsweise im naturwissenschaftlichen Unterricht mit Methoden naturwissenschaftlichen Arbeitens vertraut gemacht werden oder lediglich Frontalunterricht mit Tafel und Kreide kennen gelernt haben? Ist es wünschenswert, dass Lehrer auf die zeitraubende Selbsttätigkeit der Schüler, auf Gruppenarbeit oder kreative individuelle Schülerleistungen verzichten, weil sich dergleichen wahrscheinlich nicht in hohen Testleistungen niederschlägt?

4. Forschungsergebnisse zur Wirksamkeit institutioneller Rahmenbedingungen

Die wenig ergiebigen Befunde zur Wirksamkeit schulischer Ressourcen sind in der bildungsökonomischen Forschung in letzter Zeit zum Anlass genommen worden, das Augenmerk stärker auf andere Strategien der Qualitäts- und Effizienzverbesserung zu richten: die verhaltenssteuernden institutionellen Rahmenbedingungen des Schulsystems.

„Aus ökonomischer Sicht versprechen solche institutionellen Rahmenbedingungen den größten Erfolg, die für alle Beteiligten Anreize schaffen, die Lernleistungen der Schüler zu erhöhen: [...] Regelungen und Regulierungen des Schulsystems, die explizite oder implizite Belohnungen und Sanktionen für unterschiedliches Verhalten der Akteure erzeugen“ (Wößmann 2005, S. 19).

Dezentralisierung, Autonomie, outputorientierte Steuerung, extern gesetzte und überprüfte Standards sowie Wettbewerbselemente gelten als leistungsfördernde institutionelle Rahmenbedingungen. Die empirische Untersuchung ihrer Wirksamkeit wird durch den Zugang zu umfangreichen Datensätzen aus internationalen Schulleistungsstudien begünstigt. Die Mühen eigenständiger Instrumententwicklung und Datengenerierung bleiben den Bildungsökonominnen damit erspart. Da die meisten institutionellen Regelungen innerhalb von Ländern normalerweise nicht variieren, wird die Möglichkeit, Institutionen auf der Basis von Daten aus ländervergleichenden Studien zu untersuchen, als einmalige Chance zur Analyse ihrer Wirkungen auf die Schülerleistungen begriffen. Das heißt, die Welt wird noch einmal als *educational laboratory* betrachtet, wie in den Zeiten der ersten IEA-Untersuchungen (Husén 1967, S. 27). Dass etliche der an den IEA-Studien beteiligten Wissenschaftler

schon früh die Hoffnung aufgegeben haben, mithilfe von Ländervergleichen Kausalfaktoren guter Schulleistungen identifizieren zu können, sei nur am Rande erwähnt (vgl. Schümer 2004).

Der Vorteil einer für das Auffinden von Effekten hinreichenden Varianz institutioneller Faktoren, den internationale Datensätze bieten, wird erkauft mit den bekannten Problemen des internationalen Vergleichs. Problemverschärfend wirkt sich aus, dass mit der Schätzung internationaler Bildungsproduktionsfunktionen über alle Länder hinweg ein Durchschnittseffekt (Meta-Effekt) für institutionelle Einflussfaktoren ermittelt wird und mit diesem Ergebnis *nationale* Politikempfehlungen begründet werden. Helmut Fend hat in seinen Arbeiten immer wieder auf das Problem einer dekontextualisierten Interpretation von Variablen hingewiesen. Einzelne institutionelle Merkmale sind Bestandteil eines komplexen – länderspezifischen - Geflechts von Bedingungsvariablen, die die Wirkung dieser Merkmale verstärken oder kompensieren können.

Übersicht 1: Schulische Faktoren und ihr Einfluss auf die Lesekompetenz in verschiedenen Ländern (Zuwachs an Punkten bei Veränderung der Prädiktoren um eine Standardabweichung)

Länder	Schulklima ¹	Arbeitshaltung und Stimmung der Lehrer ¹	Lehrer-autonomie ²	Schul-autonomie ²	Schüler-Lehrer-Verhältnis ³	Schuldisziplin ³	Leistungsdruck ³
Australien	8	7	-6	14	8	19	-6
Kanada	4	2	2	18	18	10	-3
Finnland	-3	8	6	-4	11	6	-13
Deutschland	4	1	-11	-8	32	1	0
Schweiz	-8	4	-2	9	11	10	-11
Vereinigtes Königreich	17	5	-2	18	19	15	-8

¹ Einschätzung des Engagements des Lehrkörpers durch Schulleiter.

² Summierung von Entscheidungsmöglichkeiten aus den Angaben der Schulleitung.

³ nach Einschätzung der Schüler.

Quelle: Fend 2004, S. 32

Die Zusammenstellung einiger Schulvariablen zeigt, dass die gleichen Variablen in verschiedenen Ländern eine ganz unterschiedliche Bedeutung haben. Fend schlägt im Blick darauf vor, statt einzelner Variablen Konfigurationen von Merkmalen¹⁰ erfolgreicher Länder in den Blick zu nehmen. Dabei wird deutlich, dass es keine allgemeingültige Konfiguration gibt, dass Erfolg mit unterschiedlichen, funktional äquivalenten Konfigurationen erreichbar ist, mit anderen Worten: dass verschiedene Wege nach Rom führen (vgl. dazu im Einzelnen Fend 2004).

4.1 Zentrale Abschlussprüfungen

Zu den Institutionen, die von Bildungsökonomern gern untersucht werden, gehören u.a. zentrale Abschlussprüfungen. Ihre Effekte sind mehrfach auf der Basis der Daten der IEA und

der OECD analysiert worden, so von Fuchs und Wößmann (siehe Fuchs & Wößmann 2007, Wößmann 2002, 2003, 2005d-f und 2006) und von Büchel, Jürges und Schneider (2003, 2004). Ausgangspunkt all dieser Studien ist die Überlegung, dass Schulsysteme u.a. dann erfolgreich sind, wenn sie „für alle Beteiligten Anreize schaffen, die Lernleistungen der Schüler zu erhöhen“, indem sie „explizite oder implizite Belohnungen und Sanktionen für unterschiedliches Verhalten der Akteure erzeugen“ (Wößmann 2005f, S. 19; vgl. die theoretischen Ausführungen in Bishop & Wößmann 2004). Aus ökonomischer Sicht gehören zentrale Prüfungen zu den besonders wirksamen Anreizen zur Verbesserung der Leistungen, da sie den Lernerifer von Schülern stärken und Lehrer veranlassen, „ihre Partikularinteressen beiseite zu schieben und sich stattdessen hauptsächlich auf die Verbesserung der Schülerleistungen zu konzentrieren“ (Wößmann 2006, S. 430).

10. Mit der Qualitative Comparative Analysis (QCA – Ragin 1987) steht eine Methode zur Identifikation solcher Konfigurationen zur Verfügung. Diese Methode, die mit dichotomisierten Variablen arbeitet, eignet sich für Analysen in Situationen, die durch wenige Fälle und viele erklärenden Variablen gekennzeichnet sind. Empirische Erfahrungen damit im Bildungsbereich konnten etwa in einer Untersuchung von Einflussfaktoren kantonaler Unterschiede in den Mathematikleistungen Schweizer Schüler in PISA 2003 gemacht werden (Holzer 2005).

Die Bedeutung von Anreizen für das menschliche Handeln ist nicht zu bestreiten. Auch der Hinweis darauf, dass Belohnungen und Bestrafungen zu den wirksamen Verhaltensanreizen gehören und dass es explizite und implizite positive bzw. negative Sanktionen gibt, ist zu akzeptieren. Man fragt sich aber vergebens, weshalb im Fall der Schule ausschließlich extrinsische Motivationen eine Rolle spielen sollen. Sicher trifft man immer wieder Schüler, die faul, und Lehrer, die bequem sind, doch sollte man darüber nicht vergessen, dass es auch intrinsisch motivierte Schüler und engagierte Lehrer gibt, die nicht nur dann etwas für die Schule tun, wenn sie für ihren Eifer bzw. ihren Arbeitseinsatz extern belohnt oder bestraft werden. Das heißt, die Voraussetzungen, die in den zitierten ökonomischen Studien gemacht werden, sind so nicht haltbar. Sie sollen hier auch gar nicht weiter diskutiert werden. Stattdessen wird gefragt, wie die Ergebnisse der Studien gewonnen wurden, wie belastbar sie sind und wie gut sich bildungspolitische Empfehlungen mit ihnen begründen lassen.

Büchel, Jürges & Schneider (2004) analysieren die deutschen TIMSS-Daten und fragen, ob sich zwischen Bundesländern mit und ohne zentrale Schulabschlussprüfungen am Ende der Pflichtschulzeit Unterschiede in den Schülerleistungen zeigen. Ihre Untersuchung beschränkt sich auf die Leistungen von Haupt- bzw. Realschülern, da es in den anderen Schulformen am Ende der Sekundarstufe I keine Prüfungen gibt. Ihre Analysen beziehen sich im Wesentlichen auch nur auf Mathematikleistungen, da in TIMSS nur mathematische und naturwissenschaftliche Leistungen getestet wurden, die naturwissenschaftlichen Leistungen aber nur für einen Teil der Schüler in Bayern und Sachsen Gegenstand von Abschlussprüfungen sind.

Diese Situation wird von den Autoren für die Schätzung von „Differenzen-in-Differenzen“ genutzt, d.h. sie prüfen, ob die individuellen Unterschiede zwischen mathematischen und naturwissenschaftlichen Leistungen in Ländern mit Prüfungen größer sind als in Ländern ohne Prüfungen. Der Effekt der Schülerzusammensetzung, d.h. dass die Länder mit Prüfungen – Bayern, Baden-Württemberg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt und Thüringen – auch Länder mit hohen Anteilen an Haupt- und Realschülern sind, wird statistisch kontrolliert. Das Problem, dass sich die Länder mit Prüfungen eventuell auch in etlichen nicht erfassten, aber leistungsrelevanten Merkmalen von den Ländern ohne Prüfungen unterscheiden, konnte nicht mit statistischen Mitteln gelöst werden. Dieses Problem ist aber keineswegs trivial: Da mit Individualdaten der Schüler gearbeitet wird, und da die Zahl der Schüler in Bayern und Baden-Württemberg mehr als doppelt so groß ist wie die Zahl der Schüler aus den vier anderen Bundesländern zusammen genommen, sagen die Ergebnisse mehr über die „Musterländer“ Bayern und Baden-Württemberg aus als über die Wirkungen zentraler Prüfungen. Immerhin beträgt der Effekt der Prüfungen nach den entsprechenden Kontrollen noch rund ein Zehntel einer Standardabweichung. Die Autoren selbst vermuten, dass dieser Leistungsvorsprung der Schüler aus Ländern mit Prüfungen „zu einem erheblichen Teil (wenn auch nicht vollständig) auf unbeobachtete Unterschiede zwischen den Bundesländern zurückzuführen“ ist (Büchel, Jürges & Schneider 2004, S. 71). Trotzdem erwarten sie, dass zentrale Prüfungen generell zu höheren Schulleistungen führen und empfehlen ihre Einführung in allen Bundesländern.

Wesentlich problematischer als die auf Deutschland beschränkten Studien von Büchel u.a. sind die Untersuchungen, die auf internationalen Daten beruhen. Sie kommen im Prinzip alle zu dem gleichen Ergebnis, nämlich, dass Schüler in Ländern mit externen Abschlussprüfungen in den Leistungstests wesentlich besser abschneiden als Schüler in Ländern ohne solche Prüfungen. Nach Befunden von Wößmann, dessen erste Arbeiten¹¹ zum Effekt zentraler Prüfungen auf Daten aus TIMSS und TIMSS-Repeat basieren, kann der Effekt zentraler Prüfungen auf die Mathematikleistungen sogar größer sein als ein Jahrgangsstufenäquivalent, d.h. größer als der durchschnittliche Lernzuwachs von Schülern im Verlauf eines ganzen Schuljahrs (Wößmann 2002, S. 20). Das waren in TIMSS rund 40 Punkte. Die Effekte sind im Fall der naturwissenschaftlichen Leistungen nicht wesentlich schwächer, und sie verschwinden in keinem der beiden Leistungsbereiche, wenn Merkmale der Schüler und ihrer Schulen in die Analysen einbezogen werden.¹² Schon aufgrund der Stärke der Effekte liegt die Vermutung nahe, dass man es auch hier mit Wirkungen nicht erfasster Ländermerkmale zu tun hat, die mit der Einrichtung von zentralen Prüfungen konfundiert sind. Wößmann glaubt, den Einfluss solcher Merkmale kontrollieren zu können, indem er Daten zu anderen, ebenfalls zentralen Institutionen in den Ländern berücksichtigt und Dummy-Variablen für Gruppen von Ländern mit vergleichbarer Kultur in die Analysen einführt.

Wenn man sich die Liste der Länder mit und ohne zentrale Prüfungen genauer anschaut, gewinnt man allerdings den Eindruck, dass diese Berechnungen überflüssig sind. Zu den Ländern ohne Prüfungen bzw. mit Prüfungen in Mathematik für weniger als die Hälfte der Schüler gehören – sieht man von Belgien, Deutschland, Österreich, der Schweiz und den USA ab – fünf südeuropäische Länder, sowie Kuwait, Chile, Kolumbien und die Philippinen. In den übrigen 40 an TIMSS oder TIMSS-Repeat beteiligten Ländern gibt es zentrale Prüfungen. Entwicklungs- oder Schwellenländer finden sich nur höchst vereinzelt unter ihnen. Das Problem nicht erfasster Einflussgrößen ist aber nicht das einzige ungelöste Problem. Ebenso schwer wiegt wohl, dass Wößmann in dieser und den folgenden Studien von der fragwürdigen Voraussetzung ausgeht, dass der Leistungsvorsprung der untersuchten Siebt- und Achtklässler eine Wirkung der in ihren Ländern institutionalisierten zentralen Abschlussprüfungen am Ende der Sekundarstufe II ist, d.h. eine Wirkung von Prüfungen, die mindestens vier, in der Regel aber fünf oder mehr Jahre später stattfinden und in vielen Ländern längst nicht alle Schüler betreffen. Schon aufgrund der von Büchel u.a. geschätzten Differenzen-in-Differenzen ist man hier skeptisch. Zudem erfährt man nichts über den Charakter der Prüfungen. Auf diesen Mangel weist Wößmann selbst hin und führt auch auf, in welcher Hinsicht sich Prüfungen voneinander unterscheiden können (a.a.O., S. 17): In manchen Ländern müssen die Schüler Abhandlungen schreiben, in anderen Multiple-Choice-Aufgaben lösen. In manchen Ländern

11. Das Entstehungsjahr ist in der Regel nicht gleichzusetzen mit dem Erscheinungsjahr der Veröffentlichungen.

12. Die Ergebnisse der Analyse der TIMSS-Daten werden durch die Ergebnisse zu TIMSS-Repeat bestätigt. Es werden auch Ergebnisse von Berechnungen auf der Basis der gepoolten Daten beider Studien präsentiert. Hier wüsste man gern mehr über die Vorgehensweise, denn die Testleistungen der Länder, die an beiden Studien teilnahmen, liegen sowohl in Mathematik als auch in den Naturwissenschaften um bis zu 55 Punkte auseinander. Viele Länder haben in TIMSS-Repeat deutlich schlechter abgeschnitten als in TIMSS.

werden die Prüfungsergebnisse der individuellen Schüler bzw. der einzelnen Schulen veröffentlicht, in anderen nicht. Zum Teil sind die Prüfungen versetzungsrelevant, zum Teil haben sie keinerlei Konsequenzen für die weitere Schullaufbahn oder den Einstieg in einen Beruf. Man kann daraus nur den Schluss ziehen, dass Analysen von Daten aus Ländern, deren Entwicklungsstand und deren Lernkultur nicht vergleichbar sind, und die höchst unterschiedliche Prüfungen institutionalisiert haben, wohl kaum generelle Aussagen über die Leistungsrelevanz von zentralen Abschlussprüfungen erlauben.

Auf Daten von PISA 2000 gestützt haben Fuchs und Wößmann ein weiteres Mal die Wirkungen von zentralen Abschlussprüfungen auf die Schülerleistungen untersucht, worunter in PISA Leistungen im Lesen, der Mathematik und den Naturwissenschaften zu verstehen sind. Die Effekte sind jetzt wesentlich schwächer¹³: In Mathematik betragen sie noch rund 19 Punkte, in den Naturwissenschaften rund 15 und im Lesen nur noch rund 7 Punkte; der Effekt der externen Prüfungen auf die Leseleistungen ist nicht signifikant. Die Wirkungen von mindestens einmal im Jahr durchgeführten standardisierten Tests sind lange nicht so stark: Sie machen im Durchschnitt nur zwischen zwei und vier Punkten aus. Berücksichtigt man allerdings Wechselwirkungen von standardisierten Tests und externen Prüfungen, kommt man zu anderen Ergebnissen: In Ländern mit beiden Formen der Leistungskontrolle sind die Leistungen in Mathematik um rund 12 Punkte besser, in den Naturwissenschaften um rund 16

Punkte und im Lesen um rund 14 Punkte. Entsprechendes gilt, wenn Interaktionen von Prüfungen und Entscheidungsbefugnissen von Schulen oder Wettbewerbselemente im Schulsystem berechnet werden. Von den Autoren selbst wird vermerkt, dass keines der Ergebnisse

Man kann daraus nur den Schluss ziehen, dass Analysen von Daten aus Ländern, deren Entwicklungsstand und deren Lernkultur nicht vergleichbar sind, und die höchst unterschiedliche Prüfungen institutionalisiert haben, wohl kaum generelle Aussagen über die Leistungsrelevanz von zentralen Abschlussprüfungen erlauben.



kausal interpretiert werden darf, da mit Wirkungen nicht erfasster Einflussgrößen zu rechnen ist. Das hindert sie aber nicht daran festzustellen, dass institutionelle Regelungen erheblich zur Aufklärung von Leistungsunterschieden zwischen Ländern beitragen.

13. Wahrscheinlich wären die Effekte noch schwächer, wenn die Autoren sich auf die OECD-Länder beschränkt hätten, anstatt auch vier kooperative Länder zu berücksichtigen, die in den von der OECD vorgelegten Analysen gesondert behandelt werden, da sie nicht mit den Mitgliedsstaaten der OECD vergleichbar sind. Wer diese Regel missachtet, muss mit Verzerrungen seiner Ergebnisse rechnen, da sich Entwicklungs- und Schwellenländer in vielfacher Hinsicht von hoch entwickelten Industrienationen unterscheiden.

4.2 Schulautonomie

In zentralen Prüfungen sehen BildungsökonomInnen eine *conditio sine qua non* für einen leistungsfördernden Effekt größerer Handlungsfreiheit der Schulen. „Damit eine Dezentralisierung des Schulsystems [...] zu positiven Leistungseffekten führt, muss durch zentrale Prüfungen sicher gestellt sein, dass die dezentralen Entscheidungsträger Anreize zu leistungsförderndem Verhalten haben“ (Gundlach & Wößmann 2003, S.31). Fehle diese Voraussetzung, könne erweiterte Handlungskompetenz der Schulen sogar kontraproduktiv sein, indem sie dazu verleite, dass Schulen eigene Interessen verfolgen. Um einen empirischen Nachweis dieser These hat sich vor allem Ludger Wößmann bemüht. In seinen Analysen von TIMSS- und PISA-Daten kann er zeigen, dass sich die Effekte von Schulautonomie (Effektstärke und Effektrichtung) nicht nur nach Entscheidungsbereichen unterscheiden, sondern auch danach, ob die Schulen durch externe Leistungsprüfungen zur Rechenschaft gezogen werden oder nicht (Wößmann 2003, 2005a, 2005b)¹⁴. Seine Analysen von PISA-Daten vermitteln u.a. folgendes Bild (Wößmann 2005 a):

In Systemen ohne zentrale Prüfungen sind die Schülerleistungen im Durchschnitt niedriger, wenn die Schulen selbstständig über Lehrinhalte und Lehrergehälter entscheiden können; in Systemen mit zentralen Prüfungen dreht sich dieser Effekt komplett um, der Zusammenhang wird positiv.

Fehlen zentrale Prüfungen, zeigt sich kein signifikanter Autonomieeffekt bei der Lehrbuchauswahl, und die Leistungen von Schülern in Schulen, die selbstständiger über Lehrinhalte entscheiden können, fallen signifikant schlechter aus; unter den Bedingungen zentraler Prüfungen zeigt sich dagegen in beiden Fällen im Durchschnitt ein substanzieller und statistisch signifikanter Effekt. In Schulsystemen mit zentralen Prüfungen hängen die Schülerleistungen wesentlich enger mit der Schulautonomie hinsichtlich der internen Mittelverwendung zusammen als in Systemen ohne zentrale Prüfungen; der Zusammenhang ist statistisch signifikant.

Wößmann erklärt diese Ergebnisse damit, dass das Fehlen externer Standards opportunistisches Verhalten der Akteure begünstige. Die Verpflichtung auf solche Standards und deren Kontrolle Sorge dagegen dafür, dass die Entscheidungsträger in den Schulen ihre Autonomie nicht zu ihrem eigenen Vorteil nutzen, sondern ihren Wissensvorsprung auf Grund der größeren Vertrautheit mit den lokalen Gegebenheiten zur Leistungsförderung der Schüler einsetzen. Dies sei umso eher zu erwarten, je stärker der erbrachte Leistungsnachweis an Belohnungen und Sanktionen gekoppelt sei, wie dies Modelle einer leistungsbezogenen Mittelzuweisung an die Schulen vorsehen.¹⁵

Eine effiziente Bildungspolitik, so das politische Resümee Wößmanns, „sollte Zentralprüfungen mit Schulautonomie verbinden, sie sollte Standards extern vorgeben und überprü-

14. Technisch erfolgt dies durch Ergänzung des Schätzmodells um einen Vektor multiplikativ verknüpfter institutioneller Variablen (siehe S. 4); dadurch wird es möglich, Interaktionseffekte zwischen externen Abschlussprüfungen und Schulautonomie zu untersuchen.

15. Eine solche leistungsabhängige Mittelzuweisung an Schulen kann entweder direkt erfolgen, auf der Basis erreichter Leistungsniveaus (payment by result) oder indirekt, über die Zahl der – im Wettbewerb mit anderen Schulen – rekrutierten Schüler. Im letzten Fall spricht man von nachfrageorientierter Finanzierung (demand-led financing). Unterstellt wird dabei, dass sich die Nachfrage an publizierten Leistungsindikatoren der Schulen orientiert.

fen und es gleichzeitig den Schulen überlassen, wie sie diese Standards erreichen wollen“. (Wößmann 2005c, S 166; Hervorhebung der Verfasser).

4.3 „Risiken und Nebenwirkungen“ verhaltenssteuernder Institutionen

Im Folgenden soll die Frage thematisiert werden, wie denn die Akteure auf diese institutionellen Rahmenbedingungen reagieren werden, wenn es ihnen überlassen bleibt, wie sie vorgegebene Standards erfüllen. Dabei wird, gestützt auf Erfahrungen mit *high stakes testing*¹⁶ in den USA und unter Rückgriff auf Erkenntnisse des soziologischen Neo-Institutionalismus, die These vertreten, dass die Bandbreite möglicher Verhaltensreaktionen weit größer ist, als einfache ökonomische Erklärungsansätze unterstellen, und dass sich darunter auch Verhaltensreaktionen befinden, die unerwünschte Wirkungen hervorbringen.

Um dies näher zu begründen, sei zunächst daran erinnert, dass wir es im Bildungsbereich mit spezifischen Produktionsbedingungen, mit Eigengesetzlichkeiten pädagogischer Prozesse, zu tun haben, deren Implikationen oft – vor allem von Ökonomen – nicht angemessen berücksichtigt werden. Dazu gehören eine hohe Unsicherheit bezüglich der eigenen Technologien (Technologiedefizit) und ein starker Einfluss „externer Mitproduzenten“ auf das Leistungsergebnis. Ohne Absicht der Desillusionierung sei daran erinnert, dass Intelligenz und soziale Herkunft der Schüler die mit Abstand erklärungsmächtigsten Schulleistungsdeterminanten sind. Angesicht dieser Rahmenbedingungen ist

es für Schulen oftmals Erfolg versprechender, vorgegebene Leistungsstandards, insbesondere wenn daran Belohnungen und Sanktionen geknüpft sind, über Mechanismen der Schülerselektion oder andere nicht-pädagogische Maßnahmen zu sichern. Vielfältige aktuelle Beispiele dafür liefern etwa die USA. Seit Einführung von *high stakes testing* im dortigen Schulsystem im Rahmen der durch den *No Child Left Behind Act* neu geschaffenen dezentralen Verantwortungsstrukturen häufen sich die Meldungen über solche Maßnahmen. Das – noch als harmlos zu bezeichnende – *teaching to the test* zählt dazu ebenso wie Formen des Betrugs: indem vorab an Testitems geübt wird, Hilfestellungen während des Tests gegeben werden oder falsche Antworten korrigiert werden. Auch aus anderen Ländern werden solche ungewollten Nebenwirkungen der neuen Steuerungssysteme berichtet. Aus England sind zudem Fälle bekannt, wonach schwache Schüler an Testtagen vom Unterricht ausgeschlossen wurden.

Übersehen wird meist auch, dass sich unter dem Einfluss der neuen Steuerungsstrukturen das Leistungsspektrum von Schulen sowie das Verständnis schulischer Bildung verändern. So ergab eine kürzlich durchgeführte Umfrage in amerikanischen Schulbezirken (Stark Rentner, Scott u.a. 2005), dass über 70 Prozent von ihnen eine Umverteilung der Unterrichtszeit zugunsten der Testfächer vorgenommen haben (manche Fächer werden offensichtlich gar nicht mehr unterrichtet). Aufschlussreich sind auch die Ergebnisse einer soeben erschienenen englischen Studie (Mansell 2007). Basierend auf Daten von 340 Primarschulen kommt die

16. *High stakes tests* werden so genannt, weil sie Konsequenzen für Schüler und Schulen haben, die bei diesen Tests schlecht abschneiden.

Studie zum Ergebnis, dass jeweils im Zeitraum Januar bis Mai die Schulen fast die Hälfte (44% !) ihrer Unterrichtszeit (im Durchschnitt 10 Stunden pro Woche) für *test coaching* aufgewendet hatten, d.h. für die Vorbereitung der Schüler für die zweite Teststufe (*key stage 2 test*)¹⁷. Hinzu kommen zusätzliche extracurriculare Vorbereitungszeiten im Umfang von über vier Stunden wöchentlich. Auch wurde der Unterricht in nicht testrelevanten Fächern gekürzt.

Dieses Ergebnis verweist auf die grundsätzliche Problematik, gemessene Verbesserungen bei evaluationsrelevanten Leistungen, die stets nur einen Ausschnitt aus dem gesamten Leistungsspektrum abbilden, ohne weiteres als realen Effizienzgewinn zu interpretieren (vgl. die ausführliche Darstellung von Weiß & Bellmann 2007). Hinzu kommt, dass bei näherer Betrachtung der Leistungsverteilung vielfach eine Benachteiligung besonders leistungsschwacher Schüler auffällt, ein Ergebnis, das sich daraus erklärt, dass für die Schulen ein Anreiz besteht, individuelle Förderung auf jene Gruppen zu konzentrieren, deren Leistungen knapp unterhalb zu erreichender Mindeststandards liegen (vgl. Neal & Schanzenbach 2007).

Aufschlussreich sind in diesem Zusammenhang auch die Erkenntnisse der neoinstitutionalistischen Organisationsforschung, die Nachahmungs- und Strukturangleichungsprozesse (mimetischen Isomorphismus) prognostizieren, wenn – wie im Bildungsbereich – unter Bedingungen technologischer Unsicherheit extern gesetzte Standards erfüllt werden müssen. Die bei uns nach dem PISA-Schock einsetzende

hektische Suche nach nachahmenswerten Vorbildern ist dafür ein augenfälliges Beispiel. Nun ist gegen mimetischen Isomorphismus im Prinzip nichts einzuwenden, wenn dies zur Durchsetzung von *best practice* auf breiter Basis führte. Abgesehen davon, dass Unterschiedlichkeit der Kontexte und vielfach fehlender Konsens über *best practice* diese Strategie an Grenzen stoßen lässt, sind auch hier unerwünschte Nebenwirkungen zu erwarten, nämlich reduzierte Innovationsdynamik und Vielfalt.

Hinweise darauf geben Studien über *Charter Schools* in den USA. Diese Schulen sind von direkten staatlichen Regelungen weitgehend entlastet und stehen in einem freien Wettbewerb um Schüler, deren Schulbesuch öffentlich finanziert wird. Aufgrund ihrer Arbeitsbedingungen würde man einen hohen Grad an Innovation und Diversifizierung erwarten. Eine Sekundäranalyse von 190 Studien zu *Charter Schools* konnte dies jedoch nicht bestätigen (Lubienski 2003). Es lassen sich zwar einige Neuerungen auf administrativer Ebene feststellen (z.B. ein flexiblerer Einsatz von Ressourcen, neue Entlohnungssysteme, eine größere Beteiligung von Eltern an administrativen Entscheidungen, Werbung); aber auf der Ebene des Unterrichts zeichnen sich *Charter Schools* trotz ihres erheblich größeren Gestaltungsspielraums häufig durch eher traditionelle Unterrichtsmethoden und Lehrpläne aus (ähnlich Murphy 2003). Damit stellt sich die grundsätzliche Frage, in welchem Maße neue Steuerungsinstrumente überhaupt bis auf die Ebene des Unterrichts durchgreifen.

17. Getestet werden hier 11-Jährige (sechste Klassen) in den Fächern Englisch, Mathematik und Naturwissenschaften.

Dies lenkt den Blick auf einen zweiten Effekt des mimetischen Isomorphismus: die Tendenz, zur Sicherung legitimatorischer und materieller Ressourcen den Umwelterwartungen durch Anpassung der *Formalstruktur* der Organisation zu entsprechen. „Die nach außen gerichtete Formalstruktur ist jedoch nur lose mit ihrer internen Aktivitätsstruktur, dem tatsächlichen Organisationshandeln, gekoppelt. Hier findet erheblich mehr *business as usual* statt, als die Beobachtung von Veränderungen auf der formalstrukturellen Ebene erwarten ließe“ (Krücken 2004, S. 13). Diese weitgehende Entkopplung von Formal- und Aktivitätsstruktur dürfte auch eine Erklärung dafür bieten, dass die zu Dezentralisierung und Schulautonomie vorliegenden empirischen Ergebnisse anderer Forschungsprogramme, die stärker auf die Frage fokussiert sind, wie die externen Steuerungsvorgaben von den Schulen verarbeitet werden, ein weniger optimistisches Bild als die bildungsökonomische Forschung vermitteln.

Die zu den Wirkungen von Dezentralisierung und Autonomie im schwedischen Schulsystem vorliegenden empirischen Forschungsergebnisse haben den Schulforscher Mats Ekholm (1997) zu der Schlussfolgerung veranlasst, dass sich die innerschulischen Arbeitsroutinen und die Organisation des Unterrichts trotz der umfassenden staatlichen Dezentralisierungsmaßnahmen und der damit verbundenen Erweiterung der Handlungsspielräume der Schulen über einen Zeitraum von fast 30 Jahren kaum verändert haben (S. 605-607). Als möglichen Grund für die hohe Stabilität der innerschulischen Arbeitskultur vermutet Ekholm, dass es wahrscheinlich „keine so große Rolle spielt, welches Steuerungsmodell man für die Schule wählt. Das innere Leben der Schule geht weiter, unabhängig von ihrem Kontroll- und Unterstützungssystem“.

Zu einem ähnlich ernüchternden Ergebnis gelangen Malen, Ogawa und Kranz (1990) in bezug auf die Wirkungen von *school based management* in den USA: „[...] the central assumption of „theories of action“ embedded in school based management lack empirical verification. There is little evidence that school based management alters influence relationships, renews school organizations, or develops the qualities of academically effective schools“. Ambivalent fallen auch die Evaluationsergebnisse zum österreichischen Schulautonomiegesetz aus (Bachmann, Iby u.a. 1996; Specht 1997). Der erhoffte Innovationsschub stellte sich nicht ein, im Wesentlichen profitieren von der erweiterten Handlungskompetenz die ohnehin aktiven Schulen. Angesichts dieser Befunde überrascht es denn auch nicht, dass die Fallstudienforschung kaum einen Zusammenhang zwischen Schulautonomie und Schülerleistungen nachweisen kann. So konstatieren Summers und Johnson (1994, S. 14) in ihrer Auswertung amerikanischer Autonomieprojekte, „that there is little evidence to support the notion that SBM [i.e. School Based Management] is effective in increasing student performance“.

Auch die Auswertung von 83 Fallstudien zum School Based Management in verschiedenen angelsächsischen Ländern durch Leithwood und Menzies (1998) liefert kaum Hinweise auf einen positiven Zusammenhang zwischen erweiterter schulischer Handlungskompetenz und Schülerleistungen: „There is virtually no firm, research-based knowledge about the direct or indirect effects of SBM on students [...] the little research-based evidence that does exist suggests that the effects on students are just as likely to be negative or positive“ (S. 340).

Nicht nur die Fallstudien-Forschung legt eine eher zurückhaltende Einschätzung der von erweiterter Handlungsautonomie zu erwartenden leistungssteigernden Effekte nahe. Auch in einer kürzlich erschienen Sekundäranalyse von Daten aus PISA 2000 (Maslowski, Scheerens & Luyten 2007) zeigt sich kein Zusammenhang zwischen verschiedenen Autonomiebereichen und Leseleistungen. Im Falle des Personalmanagements geht zwar ein höherer Grad an Autonomie mit besseren Leistungen einher; doch verschwindet dieser Zusammenhang, wenn Unterschiede in der Schülerzusammensetzung zwischen den Schulen berücksichtigt werden. Dieser Befund legt die Vermutung nahe, dass in den Produktionsfunktionsschätzungen von Wößmann Kompositionseffekte nicht angemessen kontrolliert wurden (Maslowski, Scheerens & Luyten 2007, S. 323).

Fazit: Die empirische Forschung zur Wirkung von Schulautonomie fällt insgesamt gesehen widersprüchlich aus. Das positive Bild, das die bildungsökonomische Forschung vermittelt, wird durch die Ergebnisse anderer Disziplinen relativiert. Für die Formulierung brauchbarer bildungspolitischer Empfehlungen reichen die aus statistischen Zusammenhangsanalysen gewonnen Informationen nicht aus; entscheidend sind vielmehr Kenntnisse von *Wirkungsmechanismen*, d.h. von jenen Faktoren auf den verschiedenen Hierarchieebenen, die zum Gelingen oder Misslingen von Dezentralisierung und Schulautonomie beitragen.

4.4 Wettbewerbselemente

Wie von Dezentralisierung und Autonomie, so werden auch vom Wettbewerb nachhaltige Qualitäts- und Effizienzverbesserungen im Schulbereich erwartet. Dass diese Erwartung vor allem von Ökonomen gehegt wird, überrascht nicht, ist doch die Bewältigung von Effizienzproblemen durch Wettbewerb sozusagen ein identitätsstiftendes Merkmal der Ökonomie. „[...] the benefits of competition are so well documented in other spheres of activity that it is quite inconceivable that more competition would not be beneficial for schools (Hanushek & Wößmann 2007, S.70). In einer ganzen Reihe von Ländern, insbesondere im angelsächsischen Raum, wurden Steuerungssysteme im Bildungsbereich etabliert, die Wettbewerbselemente als konstitutiven Bestandteil beinhalten, z.B. Schulwahlfreiheit, Stärkung der Konkurrenz durch private Bildungsangebote und Formen nachfrageorientierter Finanzierung der Schulen (Pro-Kopf-Zuweisungen an die Schulen, Aushändigen von Bildungsgutscheinen an die Eltern). Nach Meinung der Befürworter von Wettbewerbssteuerung im Schulbereich erzwingt das Damoklesschwert „Klientenverlust“ – und damit Ressourcenverlust – eine besondere Anstrengungsbereitschaft, leistungssteigernde Innovationen und eine produktive Ressourcenverwendung auf Seiten der schulischen Akteure, um mit hohen Leistungsstandards im Wettbewerb bestehen zu können. „The relative lack of competition in the schooling sector tends to dull the incentives to improve quality while holding down costs“ (Bishop & Wößmann 2001, S. 2).¹⁸

18. Die Politik hat sich diese Sichtweise zum Teil zu eigen gemacht. So begründete z.B. die Regierung Thatcher gegenüber den Eltern die Einführung eines nachfrageorientierten Finanzierungssystems im Schulbereich mit zu erwartenden Qualitätsverbesserungen: „Your choice of school directly affects that school's budget – every extra pupil means extra money for the school. So your right to choose will encourage schools to aim for the highest possible standard“ (zitiert nach Bowe, Gewirtz & Ball 1994, S. 67).

Sichtet man bei der Suche nach empirischer Evidenz für Wettbewerbseffekte im Schulbereich die einschlägige Forschungsliteratur, dann zeigt sich erneut ein uneinheitliches Bild. Studien, die einen positiven Wettbewerbseffekt auf Schülerleistungen ermitteln (z.B. Hoxby 1994; Dee 1998; Böhlmark & Lindahl 2007), stehen Untersuchungen gegenüber, die keinen (signifikanten) Wettbewerbseffekt nachweisen können (z.B. Levacic & Woods 2000; Buddin & Zimmer 2005; Gibbson, Machin & Silva 2006; Carnoy, Adamson u.a. 2007; vgl. auch die Meta-Analyse von Belfield & Levin 2002). Ähnlich widersprüchlich fallen die Befunde zu nachfrageorientierten Finanzierungskonzepten aus (vgl. Belfield 2001; Weiß 2007). Von überzeugten Wettbewerbsapologeten werden ihren Erwartungen widersprechende Forschungsbefunde geflissentlich ignoriert, insbesondere wenn sie aus Disziplinen stammen, die mit anderen Hypothesen und – aus der Sicht von Ökonomen inferioren – Methoden arbeiten. Unbeirrt wird der Politik weiterhin Wettbewerb als aussichtsreiche Strategie der Qualitäts- und Effizienzverbesserung im Schulbereich angeeignet. Als besonders wirkungsvoll wird dabei die Stärkung privater Bildungsangebote erachtet. Nicht nur seien Schulen in privater Trägerschaft staatlichen Schulen leistungsmäßig überlegen, sondern der von ihnen ausgehende Konkurrenzdruck auf die staatlichen Schulen Sorge bei diesen auch für eine Leistungssteigerung. Zudem werde mit der Erweiterung der Wahlmöglichkeiten eine bessere Abstimmung von Bildungsangebot und Elternpräferenzen erreicht, wodurch die alloкатive Effizienz im Bildungsbereich verbessert werde.

Leistungssteigerung durch Privatschulen?

Wößmann (2003, 2005b) gelangt in Schätzungen einer internationalen Bildungsproduktionsfunktion auf der Basis von TIMSS-Daten zu dem Ergebnis, dass Schüler in Ländern mit einem höheren Anteil privat geleiteter Schulen im Durchschnitt bessere Leistungen aufweisen. Ein um 25 Prozentpunkte höherer Anteil privat geleiteter Schulen geht im Durchschnitt mit einem Leistungsvorteil von 13,5 Punkten in Naturwissenschaften und 15 Punkten in Mathematik einher. Wößmann sieht darin die These bestätigt, dass sich Wettbewerb durch Privatschulen positiv auf das gesamte Schulsystem auswirke (Wößmann 2005b, S 24).

Wie gelangt Wößmann zu dieser Schlussfolgerung? Ermittelt wird über alle Länder hinweg ein Durchschnittseffekt unter Außerachtlassung jeglicher Länderspezifika des Privatschulsektors. Privatschulen höchst unterschiedlicher Ausrichtung und funktionaler Bedeutung innerhalb des Schulwesens eines Landes werden in einen Topf geworfen: private Eliteschulen, Schulen mit religiöser oder besonderer pädagogischer Prägung, öffentlich finanzierte und kontrollierte sowie vom Staat unabhängige Schulen. Durchschnittseffekte lassen keine Rückschlüsse auf einzelne Länder zu und sind damit für nationale Politikempfehlungen unbrauchbar.

Der PISA-Datensatz erlaubt insofern eine differenziertere Betrachtung als der TIMSS-Datensatz, als er auch Informationen über die Höhe des öffentlichen Finanzierungsanteils der Stichprobenschulen in privater Trägerschaft enthält. Wößmann nutzt diese Information und vergleicht den Leistungsstand der 15-Jährigen zwischen vier Gruppen von PISA-Teil-

nehmerländern, die sich in der Höhe des Privatschüleranteils und des öffentlichen Finanzierungsanteils unterscheiden (hoch vs. niedrig). Danach schneiden Schüler in den Ländern mit Abstand am besten ab, die einen hohen Anteil von Schulen in privater Trägerschaft mit einem hohen Anteil von Schulen verbinden, die öffentlich finanziert werden (Wößmann 2005b, S. 24). Wößmann wertet dies wiederum als Beleg für einen Wettbewerbseffekt, denn öffentliche Finanzierung erlaube mehr Eltern die Wahl privater Schulen; dadurch entstünde mehr Wettbewerb mit Anreizen zur Leistungssteigerung (Wößmann, 2005b, S. 25). Zu sehen ist dabei, dass nur drei PISA-Teilnehmerstaaten – Belgien, Irland und die Niederlande (die übrigens in der ersten PISA-Studie auf Grund der Nichterfüllung der vorgegebenen Ausschöpfungsquote in der Datenauswertung ausgeschlossen wurden) die Bedingung eines hohen Privatschüleranteils und eines hohen Anteils an öffentlich finanzierten Privatschulen erfüllen.

Nicht nur die schmale empirische Basis lässt Zweifel an der Belastbarkeit dieses Ergebnisses aufkommen, sondern auch die eingeschränkte Validität des Privatschüleranteils eines Landes als Wettbewerbsindikator. Schulische Bildung ist eine ortsgebundene personale Dienstleistung. Für das Anbieter- und Nachfragerverhalten ist entscheidend, welche *lokalen* Wettbewerbsbedingungen jeweils vorliegen und wie diese Bedingungen von den Akteuren perzipiert werden. Auf die Problematik struktureller Wettbewerbsindikatoren verweisen etwa Levacic/Woods (2000) in ihrer Studie zur Leistungswirksamkeit von „Quasi-Märkten“ im englischen Sekundarschulwesen.

Zwischen subjektiven Wettbewerbsindikatoren (die von den Schulleitungen wahrgenommene

Wettbewerbsintensität) und strukturellen Wettbewerbsindikatoren zeigt sich nur ein schwacher oder überhaupt kein Zusammenhang (S. 68). Übersehen wird meist auch, dass private und staatliche Schulen oftmals gar nicht auf einem einheitlichen Markt, sondern auf segmentierten Märkten Bildungsleistungen – zur Befriedigung differenter Bedürfnisse – anbieten (vgl. z. B. Davies/Quirke 2007). Englische Fallstudien (Davies/Adnet/Mangan 2002) zeigen zudem, dass identische objektive lokale Wettbewerbsbedingungen mit höchst unterschiedlichen Reaktionen der schulischen Akteure einhergehen können, die von Nischenmarketing über aggressive Strategien der Marktpositionierung bis hin zu kooperativen Arrangements reichen. Und: Auch das faktische Verhalten der Bildungsnachfrager weicht mehr oder weniger stark von den Verhaltensprämissen des Marktmodells ab. Zu den empirisch vielfach belegten Auffälligkeiten zählt z.B., dass Eltern von ihrem Recht auf freie Schulwahl häufig keinen Gebrauch machen, die Exit-Option bei unzureichenden Leistungswerten einer Schule von ihnen oftmals nicht wahrgenommen wird und Schulwahlentscheidungen in vielen Fällen nach pragmatischen Gesichtspunkten getroffen werden (z.B. Wohnortnähe der Schule).

Sind Privatschulen leistungsfähiger?

Ein Ausbau privater Bildungsangebote wird nicht nur unter Hinweis auf den davon erwarteten leistungsfördernden Wettbewerb im gesamten Schulsystem propagiert, sondern vielfach auch mit dem Argument, Privatschulen seien staatlichen Schulen leistungsmäßig überlegen. Wie kürzlich im Wirtschaftsmagazin Capital zu lesen war, ist dies auch die Meinung der weit überwiegenden Mehrheit (72%) einer

vom Institut für Demoskopie Allensbach (IFD) befragten „Entscheider-Elite“.

Wird dies durch empirische Evidenz gestützt? Auch dazu sei zunächst wiederum Wößmann (2005b) zitiert. In einer mit der bekannten Methodik vorgenommenen Auswertung des internationalen Datensatzes aus der ersten PISA-Erhebung, der auch Informationen über die Trägerschaft der Stichprobenschulen enthält, gelangt er zu dem Ergebnis, dass „im Durchschnitt [...] Schüler an Schulen in privater Trägerschaft in allen drei PISA-Kompetenzbereichen statistisch signifikant besser [abschneiden] als Schüler an Schulen in öffentlicher Trägerschaft, auch nachdem die Einflüsse zahlreicher weiterer Faktoren wie die des familiären Hintergrunds herausgerechnet wurden“ (S. 24). Der Leistungsvorsprung macht mehr als ein halbes Jahrgangsstufenäquivalent aus. Dieser Meta-Effekt bildet die Basis für die Empfehlung einer stärkeren Privatisierung des Schulbereichs, eine Empfehlung, die sich auch der „Aktionsrat für Bildung“ und andere (z.B. das Institut der Deutschen Wirtschaft) zu eigen gemacht haben.

Betrachtet man die von der OECD vorgelegten *länderspezifischen* Ergebnisse von PISA 2003 zu den Unterschieden in den mittleren Leistungen zwischen öffentlichen und privaten Schulen, findet sich kaum noch Evidenz für eine Überlegenheit der Privatschulen (vgl. Übersicht 2). In den Ländern, in denen die 15-Jährigen der Privatschulen bessere Mathematikleistungen aufweisen, reduziert sich der vermeintliche Vorsprung der Privatschulen beträchtlich – in den meisten Ländern um etwa die Hälfte –, wenn der Einfluss des sozioökonomischen Hinter-

grunds der Schüler als *Individualmerkmal* statistisch kontrolliert wird (Typ A-Effekt). Wird darüber hinaus der Tatsache Rechnung getragen, dass ein zusätzlicher Leistungseinfluss von der *Zusammensetzung der Schülerschaft* einer Schule

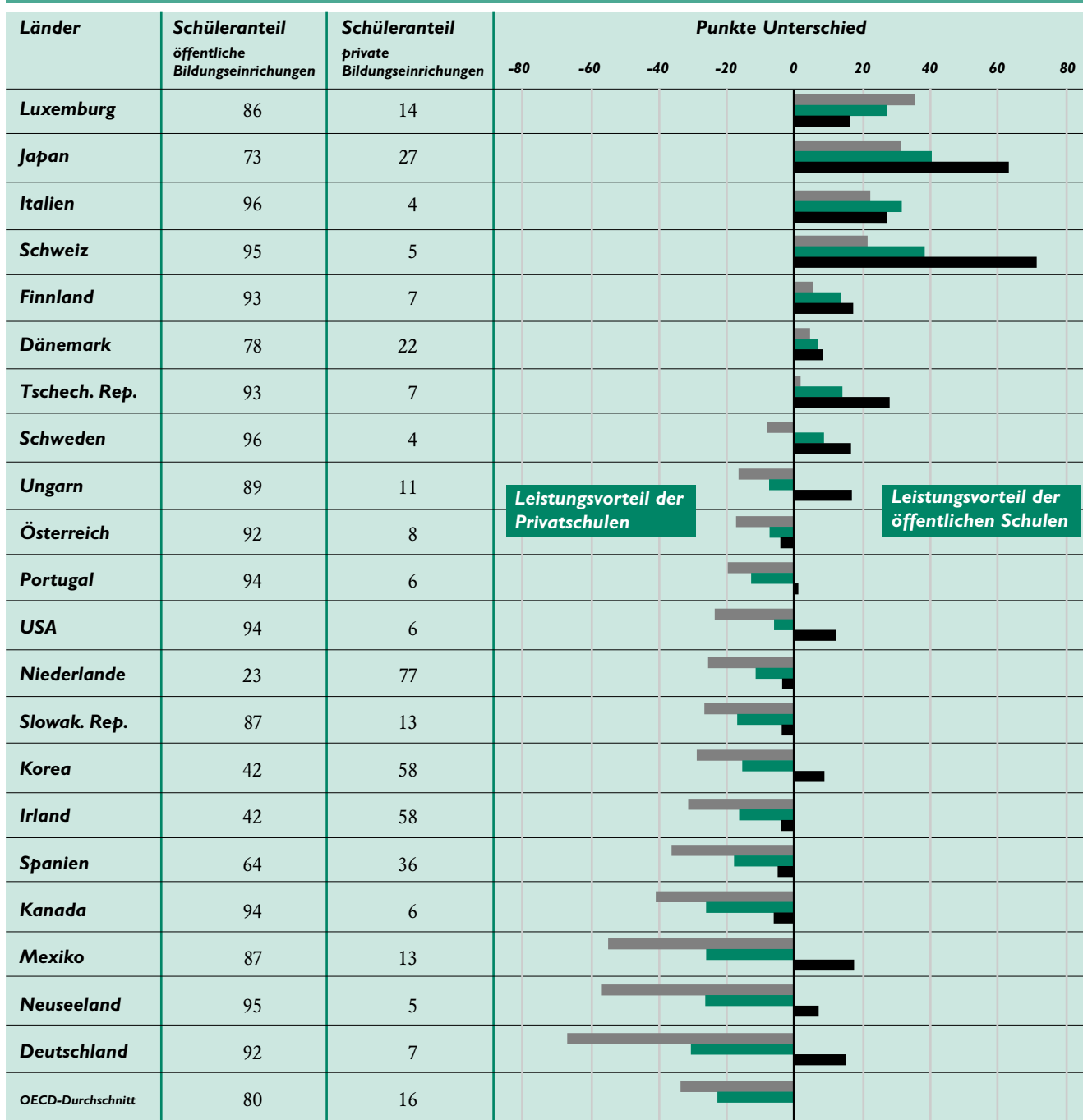
Betrachtet man die von der OECD vorgelegten länderspezifischen Ergebnisse von PISA 2003 zu den Unterschieden in den mittleren Leistungen zwischen öffentlichen und privaten Schulen, findet sich kaum noch Evidenz für eine Überlegenheit der Privatschulen.



ausgeht (Typ B-Effekt)¹⁹, ist der Leistungsvorsprung der Schüler privater Schulen entweder nicht mehr signifikant oder er verkehrt sich sogar in einen Leistungsnachteil. Das ist immerhin in sieben Ländern – darunter auch Deutschland – der Fall.

19. Die Unterscheidung zwischen Typ-A- und Typ-B-Effekt stammt von Willms & Raudenbush (1989)

Übersicht 2: Anteil der Schüler (in Prozent) und Mathematikleistungen in öffentlichen und privaten Bildungseinrichtungen (PISA 2003)



Beobachteter Leistungsunterschied
 Leistungsunterschied nach Berücksichtigung des sozioökonomischen Hintergrunds der Schüler
 Leistungsunterschied nach Berücksichtigung des sozioökonomischen Hintergrunds der Schüler und der Schule

Quelle: OECD (2004), S. 288

Schneiden Schüler in Privatschulen in PISA-E besser ab?

Einen differenzierteren Einblick in die Situation in Deutschland vermittelt ein von Weiß und Preuschoff (2006) durchgeführter Leistungsvergleich bei 15-Jährigen bzw. Neuntklässlern in staatlichen und privaten *Realschulen* und *Gymnasien*. Die Datenbasis stammt aus der Ergänzungsstudie zu PISA 2000. Unter den knapp 1.500 Schulen befinden sich 14 private Realschulen und 18 private Gymnasien. Verglichen werden die Leistungen von „statistischen Zwillingen“, das heißt von 15-Jährigen bzw. Neuntklässlern, die sich in leistungsrelevanten Merkmalen gleichen, und zwar in der sozialen Herkunft, den kognitiven Grundfähigkeiten und im Migrationshintergrund. Die Ergebnisse vermitteln ein uneinheitliches Bild mit insgesamt wenig bedeutsamen Leistungsunterschieden (vgl. Übersicht 3). Bei den Realschulen fällt der Leistungsvergleich in der Lesekompetenz und in den Naturwissenschaften zugunsten der 15-jährigen Privatschüler aus. Dieser Leistungsvorteil ist bei Differenzierung nach dem Geschlecht nur für die Mädchen nachweisbar und hängt offenbar damit zusam-

men, dass in den privaten Realschulen ein höherer Anteil der 15-Jährigen bereits die zehnte Klasse besucht. Bei den Gymnasien fällt der Leistungsvergleich insgesamt zugunsten der staatlichen Schulen aus; dort weisen die 15-Jährigen in Mathematik und in den Naturwissenschaften einen moderaten Leistungsvorsprung auf. Zusätzlich zur Leistungsuntersuchung wurden in der Studie weitere vergleichende Datenauswertungen zu verschiedenen Aspekten vorgenommen. Keine Bestätigung findet dabei die verbreitete These, Privatschulen hätten auf Grund ihrer besonderen Reformorientierung eine Vorbildfunktion für staatliche Schulen. Insgesamt zeigen sich bei pädagogischen und organisatorischen Reformkonzepten keine nennenswerten Unterschiede zwischen den staatlichen und privaten Stichprobenschulen. Vorteile weisen die Privatschulen der Stichprobe dagegen zum Teil beim Schulklima auf. Besonders ins Auge fällt dabei das Ergebnis, dass sich die 15-Jährigen besser von ihren Lehrern unterstützt fühlen als die der staatlichen Schulen. Günstiger fallen auch die Werte zur Zufriedenheit der Eltern aus, und das trotz größerer Klassen an den Privatschulen.

Übersicht 3: Leistungsunterschiede zwischen vergleichbaren Schülern staatlicher und privater Schulen

	Realschulen		Gymnasien	
	15-Jährige	Neuntklässler	15-Jährige	Neuntklässler
Lesekompetenz	6*	-2	3	5*
Mathematik	2	-7*	-5*	-5
Naturwissenschaften	11*	6	-6*	-4

Datenbasis: PISA-E 2000; Mittelwert = 500, Standardabweichung = 100;
 positives Vorzeichen = Leistungsvorsprung der Privatschulen, negatives Vorzeichen = Leistungsungleichheit der Privatschulen;
 Punktwertdifferenzen mit * sind statistisch signifikant (Weiß & Preuschoff 2006).

Neue Studien aus den USA

Auch eine neue amerikanische Vergleichstudie zu den Mathematikleistungen von Viert- und Achtklässlern kann keine Überlegenheit autonomer Schulen in privater Trägerschaft (*Charter Schools*, staatlich unabhängige Privatschulen) gegenüber staatlichen Schulen feststellen. In einer Mehrebenenanalyse von Daten des *National Assessment of Educational Progress (NAEP)* von 2003 gelangen Lubienski und Lubienski (2006) zu dem Ergebnis, dass nach Kontrolle

von Unterschieden in demographischen Schüler- und Schulmerkmalen der ursprüngliche Leistungsvorteil der privat geführten Schulen verschwindet und sich in den meisten Fällen sogar in einen Leistungsnachteil umkehrt. Auch Analysen der Leistungsentwicklung von Schülern, die staatliche Schulen und *Charter Schools* besuchten, ergaben signifikant schlechtere Ergebnisse in Lesen und Mathematik für die autonomen Schulen in privater Trägerschaft (Bifulco & Ladd 2006).

5. Institutionelle Rahmenbedingungen und Chancengleichheit

Das Modell der Bildungsproduktionsfunktion wird nicht allein dafür verwendet, die Leistungswirksamkeit von Ressourcen und Institutionen zu untersuchen, sondern ist auch Ausgangspunkt der Bearbeitung von bildungspolitisch brisanten Themen wie der Frage nach der Verwirklichung von Chancengleichheit im Bildungswesen.

Da sich das Ausmaß der Chancengleichheit daran ablesen lässt, wie stark die sozioökonomische und soziokulturelle Herkunft von Schülern ihren Schulerfolg bestimmt, lässt es sich mühelos in das Modell einordnen, d.h. als bildungsökonomisches Problem definieren: „[...] the estimates of equality of educational opportunity are derived by estimating the ‚home production‘ part of education production functions“ (vgl. Wößmann 2004, S. 2.) So gerechtfertigt wird von Wößmann bzw. von ihm und verschiedenen Koautoren (Hanushek & Wößmann 2005, Schütz & Wößmann 2005, Schütz, Ursprung & Wößmann 2005) untersucht, wie eng die Testleistungen der Schüler der einzelnen Länder mit ihrer familialen Herkunft zusammenhängen. Dies geschieht auf der Basis von Daten aus den jüngsten internationalen Schulleistungstudien: TIMSS 1995, 1999 und 2003, PISA 2000 und 2003 sowie PIRLS 2001.

Die meisten Ergebnisse der eben erwähnten ökonomischen Arbeiten sind nicht neu. Jedenfalls ist seit dem Erscheinen der Untersuchungsberichte zu den PISA-Studien und zu PIRLS sehr wohl bekannt, dass die Leistungen der Schüler in *allen* untersuchten Ländern mehr oder weniger eng mit ihrem sozioökonomischen und kulturellen Hintergrund zusammenhängen, und dass Deutschland zu denjenigen OECD-Ländern gehört, in denen der Schulerfolg besonders stark von der familialen Her-

kunft der Heranwachsenden mitbestimmt wird. Da die Schulsysteme in der Mehrzahl der an den Studien beteiligten Länder anders als das deutsche Schulsystem strukturiert sind, drängt sich angesichts der Befunde die Frage auf, ob das – im Vergleich zu anderen Ländern relativ starke – Ausmaß der sozialen Disparitäten der Schülerleistungen in Deutschland durch die hier institutionalisierte Schul- und Unterrichtsorganisation mitbedingt ist. Die Frage ist im Anschluss an die verschiedenen Untersuchungsberichte schon oft gestellt und teils positiv, teils negativ beantwortet worden. Für die Art der Antwort waren in aller Regel unterschiedliche bildungspolitische Zielvorstellungen maßgebend. Wößmann und seine Koautoren sind die ersten, die den Versuch unternehmen, mithilfe statistischer Analysen der internationalen Daten festzustellen, welchen Einfluss die Schulstruktur auf die sozial bedingten Leistungsdisparitäten hat.

Die Autoren der internationalen Schulleistungsvergleiche haben mit guten Gründen auf entsprechende Versuche verzichtet: Sie kennen die Anlage der Studien, die Vergleichbarkeit der Stichproben und der Tests und sie besitzen aufgrund des kontinuierlichen, intensiven Austauschs mit Vertretern der Teilnehmerländer nicht nur profunde Kenntnisse der untersuchten Schulsysteme sondern auch der sozialen und kulturellen Kontexte, in die sie eingebettet sind. Zwar geben einige Autoren ihre Präferen-

zen für Schulsysteme ohne externe Leistungs-differenzierung in der Öffentlichkeit sehr deutlich zu erkennen, in den Ergebnisberichten sucht man aber vergebens nach Belegen dafür, dass das Niveau und die sozialen Disparitäten der Schülerleistungen in den Teilnehmerländern durch die Struktur ihrer Schulsysteme bedingt sind. Man ist daher neugierig, wie Wößmann und seine verschiedenen Koautoren mit den internationalen Daten umgehen. Um keine Missverständnisse aufkommen zu lassen, sei betont, dass es im folgenden Text *nicht* um ein Plädoyer für oder gegen bestimmte Schulstrukturen geht. Zur Debatte steht lediglich, ob die verwendeten Daten hergeben, was Wößmann und seine Koautoren sich von ihnen versprechen, ob die Analysen zu belastbaren Ergebnissen führen und ob sie die Schlussfolgerungen rechtfertigen, zu denen die Autoren kommen.

In seinem ersten Beitrag zum Thema „Chancengleichheit“ arbeitet Wößmann (2004) mit Daten der TIMS-Studie von 1995, um den Zusammenhang von Leistung und sozialer Herkunft in den Schulsystemen von 15 westeuropäischen Ländern und den USA zu untersuchen. Seine Analysen bestätigen den seit PISA 2000 bekannten Sachverhalt, dass der Zusammenhang stark zwischen den Teilnehmerländern schwankt und nichts mit ihren durchschnittlichen Testleistungen zu tun hat. Wößmann wertet dies als Hinweis darauf, dass ein höheres Maß an Chancengleichheit im Schulwesen nicht zwangsläufig zu schlechteren Schulleistungen führt. Gegen allgemeine Feststellungen dieser Art ist nicht das Geringste einzuwenden. Problematisch aber sind die Versuche, das Ausmaß der Chancengleichheit auf

wenige Merkmale der Schulsysteme zurückzuführen, wie es in den anderen Beiträgen geschehen ist.

Schütz und Wößmann (2005) haben auf der Basis von Daten aus den TIMS-Studien von 1995 und 1999 die Effekte der familialen Herkunft auf die Schülerleistungen für 29 OECD-Länder berechnet und dargestellt, dass die Effektstärken systematisch mit dem Alter der Schüler bei der ersten Selektion für verschiedene Schulformen variieren. Dabei haben sie sich freilich Ungenauigkeiten zuschulden kommen lassen. Beispielsweise wird die Türkei zu den Ländern mit Einheitsschulsystemen gerechnet, obwohl das Land in mehrfacher Hinsicht nicht mit diesen Ländern vergleichbar ist:²⁰ Bis heute beträgt die Einschulungsquote in der Türkei in der Grundschule nur 94 Prozent und in den Klassen 6 bis 8 der weiterführenden Schulen nur 80 Prozent; rund 600.000 schulpflichtige Mädchen besuchen überhaupt keine Schule. Erst 1997 – d.h. zwei Jahre nach TIMSS 1995 – wurde gesetzlich geregelt, dass die Schulpflicht künftig acht statt wie bisher fünf Jahre betragen und der Übergang auf weiterführende Schulen *nach* der achten Klasse erfolgen solle; diese Reformen sind aber noch keineswegs verwirklicht. Auf dem Land und in den Armenvierteln der Großstädte gehen die Schulpflichtigen nach wie vor allenfalls fünf Jahre zur Schule und der Übergang auf die selektiven weiterführenden Schulen (*Anatolian High School* und *Anatolian Imam and Preacher High School*) findet wie früher bereits nach der fünften Klasse statt. Wer die Eignungstests für diese Schulen nicht besteht, besucht die *Junior High School*, die dementsprechend geringes Ansehen genießt.

20. Vermutlich liegen der Einordnung der Türkei Angaben der Europäischen Kommission zugrunde, auf die der Text von Schütz, Ursprung und Woessmann (2005) verweist. Diese Angaben dürften politisch motiviert sein. Nach verschiedenen anderen Darstellungen (z.B. OECD 2004, Auswärtiges Amt 2007) entsprechen sie nicht der Realität.

Außerdem haben die Autoren zwar den Zeitpunkt der ersten Selektion der Schüler berücksichtigt, aber vollkommen außer Acht gelassen, wie stark die einzelnen Schulsysteme gegliedert sind und wie sich die Schüler der betreffenden Altersjahrgänge auf die verschiedenen Schulformen verteilen. Zum Beispiel werden – nach Daten aus der 2003 durchgeführten PISA-Studie, auf die sich Schütz und Wößmann unter anderem beziehen – in Österreich ebenso wie in Deutschland bereits die 10-Jährigen ihren Leistungen entsprechend extern differenziert. In Österreich gibt es aber nicht drei oder vier verschiedene Schularten für diese Altersgruppe sondern lediglich zwei: Rund 30 Prozent der Schüler gehen von der fünften Klasse an auf „Allgemeinbildende Höhere Schulen“; die übrigen werden – von wenigen Sonderschülern abgesehen – gemeinsam unterrichtet und müssen sich erst mit 15 Jahren für eine von vier verschiedenen Schularten entscheiden (Bundesministerium 2002). Entsprechendes gilt für die ehemaligen sozialistischen Republiken, die nach der Wende wieder Höhere Schulen eingeführt haben (Robitaille 1997): Diese wurden in den 1990er Jahren jeweils nur von 10 bis 15 Prozent der Altersgruppe besucht; bis heute bleibt die überwältigende Mehrheit der Schüler bis zum Ende der Pflichtschulzeit zusammen. Folglich liegt die Frage nahe, ob die Klassifikation von Schütz und Wößmann den Schulsystemen dieser Länder angemessen ist.

Auch die nähere Betrachtung der Einheitschulsysteme stellt den Sinn der Klassifikation in Frage, da es in und außerhalb der Schulen die verschiedensten Formen der Leistungsdifferenzierung gibt, die ähnliche Funktionen erfüllen wie vertikal gegliederte Schulsysteme.

Zu den funktionalen Äquivalenten, die man schulintern in etlichen Ländern findet, gehören:

- *tracking* oder *streaming*, d.h. die an den Schülerleistungen orientierte Zusammensetzung von Klassen,
- *setting*, d.h. die dem leistungsdifferenzierten Fachunterricht in deutschen Schulen vergleichbare, zeitlich befristete Zusammensetzung von Lerngruppen nach der Leistung der Schüler,
- Kurssysteme, die es den Heranwachsenden erlauben, ihren persönlichen Stärken und Schwächen entsprechend in den einzelnen Fächern unterschiedlich anspruchsvolle Kurse zu besuchen,
- die Bildung von Klassen nach Kriterien, die nicht unabhängig von den Leistungen der Schüler bzw. von den Bildungsaspirationen ihrer Eltern sind, wie z.B. nach der gewählten Fremdsprache oder dem Wahlpflichtfach.

Ferner gibt es in manchen Ländern indirekte Formen der externen Differenzierung der Schüler nach Leistung. Dazu gehören große sozioökonomische Unterschiede zwischen verschiedenen Schuleinzugsbereichen, wie man sie beispielsweise in den Vereinigten Staaten häufig findet, und die Verbreitung von Privatschulen oder Schulen mit besonderen Schwerpunkten bzw. attraktiven Angeboten. Schließlich sind in diesem Zusammenhang auch die selektiven kommerziellen Ergänzungsschulen zu nennen, die in einigen ostasiatischen Ländern von der Mehrheit der Sekundarschüler besucht werden.

Die fragwürdige Klassifikation der Schulsysteme weckt Zweifel daran, dass sich die Aussage, späte Selektion könne den Einfluss der familia-

len Herkunft kompensieren und zu einem Mehr an Chancengleichheit führen, durch vergleichende Analysen internationaler Daten empirisch absichern lässt. Die Zweifel werden durch die weiteren Analysen der Autoren nicht ausgeräumt, sondern eher noch verstärkt. Um den Einwand zu entkräften, sie ließen den Einfluss unbeobachteter Heterogenität der Länder unberücksichtigt, messen Schütz und Wößmann bzw. Hanushek und Wößmann die Effekte der Schulstruktur auf das Ausmaß an Chancengleichheit und das Leistungsniveau mithilfe von Differenzen-in-Differenzen-Schätzungen: Sie untersuchen die Streuung der Grundschulleistungen in den einzelnen Ländern, nutzen sie zur Vorhersage der Leistungsstreuung bei den Sekundarschülern und prüfen, ob sich die Leistungen in mehrgliedrigen Schulsystemen stärker auseinander entwickelt haben als in Einheitsschulsystemen. Gerechnet wird mit den länderspezifischen Standardabweichungen der Grundschulleistungen, wie sie in PIRLS und TIMSS (1995 und 2003) gemessen wurden, und der Sekundarschulleistungen, wie sie in PISA (2000 und 2003) und in TIMSS (1995, 1999 und 2003) erfasst wurden. Das regressionsanalytische Verfahren hält solche länderspezifischen Rahmenbedingungen konstant, die sich u.a. im Ausmaß der Streuung der Grundschul- und der Sekundarschulleistungen niederschlagen, wie z.B. hohe Anteile von Schülern aus Migrantenfamilien.

Insgesamt wurde in den beiden Studien achtmal geprüft, ob die Streuung und das Niveau der Leistungen von Sekundarschülern systematisch mit der Mehrgliedrigkeit der Schulsysteme zusammenhängen, die an den internationalen Erhebungen in Grund- und Sekundarschulen teilgenommen haben. Bei der ersten Regressionsanalyse, die a) die Standardabweichungen

und b) die Mittelwerte der Leseleistungen in PISA 2003 auf die Standardabweichungen der Leseleistungen in PIRLS und die Schulstruktur der Teilnehmerländer zurückführt, entsprechen die Ergebnisse den Erwartungen der Autoren. Das heißt, es sieht so aus, als ob die frühe Selektion der Schüler dazu führe, dass sich ihre Leseleistungen auf der Sekundarstufe stärker als in nicht-selektiven Schulsystemen auseinander entwickeln und hinter den Durchschnittsleistungen in diesen Systemen zurückblieben; beide Effekte sind signifikant. Bei der Analyse von Daten aus PISA 2000 anstelle von PISA 2003 zeigt sich ebenfalls ein signifikanter negativer Effekt der Schulstruktur auf das Niveau der Leseleistungen; die frühe Mehrgliedrigkeit hat nach den Ergebnissen dieser Analyse aber *keine* negativen Auswirkungen auf die Leistungsstreuung. Das lässt vermuten, dass die Befunde test- oder auch stichprobenabhängig sind.

Die Vermutung wird durch die Feststellung gestützt, dass sich die Variable „frühe Mehrgliedrigkeit“ signifikant auf die Standardabweichungen der Leistungen in den naturwissenschaftlichen Tests auswirkt, und zwar bei allen drei Regressionsanalysen, während sie nur bei einer der drei Analysen der Mathematikleistungen zu entsprechenden signifikanten Ergebnissen führt. Hinsichtlich des Leistungsniveaus ist das Bild uneinheitlich: Bei den Leseleistungen ergeben beide Analysen signifikante negative Effekte der Mehrgliedrigkeit. In Mathematik kommt dies nur bei einer der drei Analysen vor und bei den naturwissenschaftlichen Leistungen zeigt sich entgegen der Erwartung einmal ein signifikant positiver Effekt der Schulsysteme mit früher Aufteilung der Schüler. Aus diesen Ergebnissen den Schluss zu ziehen, „dass eine frühe Selektion der Schüler in verschiede-

ne Schultypen nicht nur die Chancenungleichheit erhöht, sondern auch das gesamte Leistungsniveau sogar eher senkt“ (Schütz & Wößmann 2005, S. 20), ist ziemlich riskant.

Mehr als bloß riskant ist es, wenn Wößmann dann die Ergebnisse des einzigen der acht Testvergleiche, der zu einer signifikanten Bestätigung beider Hypothesen führt – Mehrgliedrigkeit führt zur Auseinanderentwicklung *und* zur Verschlechterung der Leistungen – zur Popularisierung seiner Befunde und Schlussfolgerungen benutzt (siehe Wößmann 2007). Dabei wird eine Abbildung verwendet, die die von der Schulstruktur abhängigen Veränderungen der Leistungsstreuung von der Grundschule zur Sekundarschule veranschaulicht. Die Türkei und Kanada werden den Ländern mit ungliederten Schulsystemen zugerechnet. Im gleichen Artikel wird eine Liste aus einer Veröffentlichung der OECD (2004) wiedergegeben, in der die Länder, die an PISA 2003 beteiligt waren, nach dem Alter der Schüler bei der ersten Aufteilung im Schulsystem geordnet sind. Interessanterweise fehlen in Wößmanns Wiedergabe der Liste sowohl die Türkei, in der die 11-Jährigen in Schulen eines gegliederten Systems übergehen, als auch Kanada, wo dies für die 13-Jährigen gilt. Würde man die Türkei aufgrund ihrer niedrigen Schulbesuchsquote nicht in die Berechnungen einbeziehen und die Zuordnung Kanadas zur Klasse der nicht-selektiven Schulsysteme korrigieren, führten die Regressionsanalysen sicher zu anderen Ergebnissen. Die Abbildung vermittelte allerdings auch dann den Eindruck, dass sich die Schülerleistungen in den mehrgliedrigen Schulsystemen stärker auseinander entwickeln als in den Systemen ohne frühe Selektion.

Dass vertikal gegliederte Schulsysteme zu einer wachsenden Diskrepanz der Leistungen von guten und schlechten Schülern beitragen, ist plausibel. Es fehlt auch nicht an Analysen nationaler Daten, die belegen, dass selektive Schulsysteme leistungsstarke Schüler begünstigen, leistungsschwache dagegen benachteiligen und dadurch zu einer stärkeren Auseinanderentwicklung der Leistungen beitragen, als es vor der Aufteilung der Schüler auf unterschiedlich anspruchsvolle Schulen zu erwarten gewesen wäre (siehe z.B. Baumert, Stanat & Watermann 2006).

Die Ergebnisse des internationalen Vergleichs, die Wößmann vorgelegt hat, sind auf den ersten Blick vielleicht eindrucksvoller; bei näherer Betrachtung wird man sie allerdings für zufallsbedingt halten müssen:

- Ein erstes Argument dafür wurde bereits vorgebracht, nämlich die Fragwürdigkeit der dichotomen Klassifikation der Schulsysteme: Mehrgliedrige Systeme können sich, abhängig vom Zeitpunkt der Gliederung, der Zahl und Art der verschiedenen Schulformen und ihrer quantitativen Bedeutung erheblich voneinander unterscheiden. Auch die Einheitsschulsysteme bilden keine in sich homogene Gruppe, da es die verschiedensten funktionalen Äquivalente für die externe Gliederung von Schülern nach Leistung gibt.
- Zweitens ist es recht unwahrscheinlich, dass sich Einflussgrößen wie die Schulstruktur *unvermittelt* auf die Lernprozesse und Leistungen der Schüler auswirken. Ob der Verzicht auf die verschiedenen Formen der Selektion eine stärkere Auseinanderentwicklung der Schülerleistungen verhindert, ob er leistungssteigernd wirkt und dazu beiträgt, soziale Disparitäten des Schulerfolgs abzubauen, dürfte vielmehr in hohem Maße davon ab-

hängen, wie man in den Schulen mit leistungsheterogenen Lerngruppen umgeht, d.h. wie der Unterricht in diesen Gruppen gestaltet wird, wie gut schwache Schüler unterstützt werden, welche personellen und räumlichen Ressourcen dafür zur Verfügung stehen und mit was für Lernmaterialien die Schulen ausgestattet sind. Man kann annehmen, dass sich die untersuchten Länder in dieser Hinsicht nicht unerheblich voneinander unterscheiden.

- Schließlich gibt es auch Differenzen zwischen Ländern hinsichtlich sozialer und kultureller Merkmale, die in den vergleichenden Schulleistungsstudien überhaupt nicht erfasst wurden, die aber die Streuung der Leistungen nicht unbeeinflusst lassen dürften. Die Differenzen betreffen beispielsweise den Stellenwert häuslichen und schulischen Lernens, die

Bedeutung von Nachhilfeunterricht und Ergänzungsschulen, die mit Schulabschlüssen verbundenen Berufs- und Lebensperspektiven, die Einstellungen der Gesellschaft zum Lesen und generell zum Lernen oder die soziale Bewertung von Begabung und Anstrengung. Merkmale dieser Art können sich – unabhängig von der Schulstruktur – sowohl auf die Leistungen und das Ausmaß ihrer sozialen Disparitäten als auch auf die Leistungsstreuung auswirken.

Da in Wößmanns Studie alle eben genannten Einflussgrößen unbeachtet bleiben, sind seine Befunde keine stichhaltigen *Belege* dafür, dass Einheitsschulsysteme sozial benachteiligten Schülern bessere Bildungschancen bieten. Über die *Realität* ist damit freilich nichts gesagt.

Zusammenfassung

Die Bildungsökonomie erfährt gegenwärtig einen bemerkenswerten Aufschwung. Dies betrifft sowohl ihre Position innerhalb der wirtschaftswissenschaftlichen Forschung als auch ihre Funktion in der Politikberatung. Die verschärften Knappheitsprobleme im öffentlichen Sektor bieten dafür ebenso eine Erklärung wie der Bedeutungszuwachs, den Bildung als strategische Ressource für die Bewältigung wirtschaftlicher und gesellschaftlicher Probleme erfahren hat. Schubkraft verliehen haben der Bildungsökonomie die Aktivitäten der OECD im Bildungsbereich sowie die thematische Akzentverlagerung innerhalb des Forschungsprogramms der Disziplin von theoretischen Fragestellungen der Wirtschaftswissenschaften zu praktischen Steuerungsproblemen des politisch-administrativen Systems.

Im Zuge ihrer programmatischen Neuausrichtung hat die Bildungsökonomie in letzter Zeit in verstärktem Maße Themen besetzt, die traditionell in den Zuständigkeitsbereich anderer Bildungswissenschaften fallen. Besonderes Augenmerk gilt dabei der *Qualität der Schulbildung*. Ihr wird in neueren bildungsökonomischen Wirkungsstudien eine herausragende Bedeutung als Wachstumstreiber attestiert. Die Bildungsökonomie hat das theoretische und methodische Arsenal der Schulqualitätsforschung erweitert. Doch basieren ihre Forschungen häufig auf wenig belastbaren Hypothesen, d.h. Hypothesen, die der Spezifik des Objektbereichs nicht angemessen Rechnung tragen. Darin manifestieren sich nicht zuletzt *unzureichende Institutionenkenntnisse*. Ein weiteres Manko ist in der *ausgeprägten Selbstreferenzialität* der bildungsökonomischen Forschung zu sehen. Eine Auseinandersetzung mit konkurrierenden Hypothesen und konträren Forschungsergebnissen anderer Disziplinen findet in der Regel nicht

statt. Dies hängt unter anderem damit zusammen, dass die meisten Wirtschaftswissenschaftler die Bildungsökonomie nicht als Teildisziplin der Bildungsforschung begreifen und in Deutschland derzeit eine institutionelle Basis für ein produktives Zusammenwirken der Bildungsökonomie mit anderen Bildungswissenschaften fehlt.

Im Mittelpunkt der bildungsökonomischen Schulqualitätsforschung stehen Untersuchungen zur Wirksamkeit bildungspolitisch beeinflussbarer Faktoren (Ressourcen, Institutionen). Diese Arbeiten basieren vorrangig auf dem Konzept der Bildungsproduktionsfunktion. Die untersuchten Variablenzusammenhänge werden darin als Input-Output-Beziehungen modelliert; unterstellt wird ein direkter Wirkungszusammenhang zwischen abhängigen und unabhängigen Variablen. Die Annahme, dass Ressourcen und institutionelle Kontextbedingungen als distale, vom Unterrichtsgeschehen entfernte Faktoren Lernergebnisse unmittelbar beeinflussen, ist theoretisch nicht zu begründen. Der bildungsökonomische Produktionsfunktions-Ansatz wird der Komplexität schulischer Bildungsprozesse nicht gerecht. Der Aussagegehalt der damit gewonnenen Ergebnisse für bildungspolitisches Steuerungshandeln, das Informationen über Wirkungsmechanismen erfordert, ist stark eingeschränkt.

Mithilfe des Produktionsfunktionsansatzes ist in der Vergangenheit von der Bildungsökonomie vor allem der Einfluss real vorfindbarer Unterschiede in der *Ressourcenausstattung* von Schulen auf Schülerleistungen untersucht worden. Das daraus in Forschungsübersichten gezogene Resümee fällt für die Bildungspolitik ernüchternd aus: Zwischen Schulressourcen und Schülerleistungen zeigt sich kein enger

und konsistenter Zusammenhang. In der darüber geführten Debatte haben vor allem divergierende Vorstellungen von Wissenschaft und Praxis von der Wichtigkeit der *Klassengröße* als Determinante der Unterrichtsqualität für Kontroversen gesorgt.



Wenn die Rahmenbedingungen nicht berücksichtigt werden, ist es nicht weiter verwunderlich, dass sich kaum einmal signifikante Wirkungen der Klassengröße auf die Schülerleistungen feststellen lassen.

Dies beweist aber nicht, dass es sie nicht gibt.

Die bildungsökonomischen Analysen der internationalen Schulleistungsdaten kommen zum Ergebnis, dass nichts dafür spricht, dass kleine Klassen zu höheren Leistungen führen. Obschon die Analysen mit der erforderlichen methodischen Sorgfalt durchgeführt wurden, sind sie unbefriedigend, da sie von falschen Voraussetzungen ausgehen und ihre Ergebnisse zu Unrecht verallgemeinern. Sie setzen die mathematischen und naturwissenschaftlichen Leistungen von Schülern in siebten und achten Klassen mit Schülerleistungen schlechthin gleich, obwohl es – abhängig vom Unterrichtsfach und der Altersgruppe – verschiedene Arten von Leistungen gibt, und obwohl Schulen auch noch andere Aufgaben haben als die „Produktion“ von Fachleistungen, z.B. die Förderung von Lernfreude und Leistungsmotivation, die Vermittlung von Arbeitstechniken und

kommunikativen Fähigkeiten oder die Anleitung zum selbständigen Arbeiten. Ferner gehen sie von der falschen Voraussetzung aus, dass sich die Effekte von Klassenfrequenzen durch die Berücksichtigung fixer Schuleffekte von den Auswirkungen der vielen anderen Einflussgrößen isolieren lassen, die ebenfalls leistungshemmend oder -fördernd wirken (wie etwa die Ausstattung der Schulen, die didaktische und methodische Gestaltung des Unterrichts oder die Qualifikationen der Lehrer). Selbst wenn dies möglich wäre, bliebe unberücksichtigt, dass Lernprozesse nicht nur *in* sondern auch *außerhalb* der Schule stattfinden, (beispielsweise bei der Erledigung der Hausaufgaben, im Nachhilfeunterricht oder in Ergänzungsschulen), und dass Leistungen im Allgemeinen Ergebnisse *langjähriger* Lernprozesse sind. Sie können unter sehr viel günstigeren oder ungünstigeren als den in TIMSS erfassten Lernbedingungen stattgefunden haben. Schließlich ist es auch von kulturellen Rahmenbedingungen wie dem Erziehungs- und Unterrichtsstil und der Wertschätzung des Lernens in den untersuchten Ländern abhängig, was in der Schule möglich ist, und welche Klassengrößen Lehrern und Schülern zugemutet werden können. Wenn die Rahmenbedingungen nicht berücksichtigt werden, ist es nicht weiter verwunderlich, dass sich kaum einmal signifikante Wirkungen der Klassengröße auf die Schülerleistungen feststellen lassen. Dies beweist aber nicht, dass es sie nicht gibt.

Die wenig ergiebigen Befunde zur Wirksamkeit schulischer Ressourcen sind in der bildungsökonomischen Forschung in letzter Zeit zum Anlass genommen worden, das Augenmerk stärker auf andere Strategien der Qualitäts- und Effizienzverbesserung zu richten: die das Verhalten der Akteure beeinflussenden *institutio-*

nellen Rahmenbedingungen des Schulsystems. Als besonders leistungsfördernd gelten Dezentralisierung und Schulautonomie, extern gesetzte Standards und zentrale Abschlussprüfungen sowie Wettbewerb durch die Einführung von Schulwahlfreiheit und nachfrageorientierte Finanzierungskonzepte. Die empirische Untersuchung ihrer Wirksamkeit wird durch den Zugang zu Datensätzen aus internationalen Schulleistungsstudien begünstigt; sie erfüllen die Voraussetzung einer für das Auffinden von Effekten hinreichenden Varianz der institutionellen Faktoren. Dieser Vorteil wird freilich mit den bekannten Problemen von Ländervergleichen erkaufte.

Schon bei der Betrachtung der deutschen Bundesländer wird deutlich, dass der Leistungsvorsprung der Schüler aus Ländern mit *zentralen Prüfungen* zu einem ganz erheblichen Teil auf unbeobachtete Unterschiede zwischen diesen Ländern und den Bundesländern ohne zentrale Prüfungen zurückzuführen ist. Beim Vergleich der Leistungen von Schülern aus unterschiedlichen Staaten verstärkt sich der Verdacht, dass die den zentralen Prüfungen zugeschriebenen Effekte Wirkungen nicht erfasseter Einflussgrößen sind. Im Hinblick auf die großen Unterschiede im Charakter der Prüfungen, in ihrer Relevanz für die untersuchte Altersgruppe und in den Bedingungen, unter denen sie durchgeführt werden, ist gar nicht zu erwarten, dass diese zentralen Prüfungen vergleichbare Wirkungen auf die Schüler ausüben und ihre Leistungen steigern. Dies ist den Autoren zum Teil durchaus bewusst. Im Bemühen um generalisierbare Aussagen werden trotzdem über alle Länder hinweg *Durchschnittseffekte für die institutionellen Faktoren* ermittelt und damit – verschiedener Kontextunterschiede ungeachtet – nationale Politikempfehlungen begründet.

Die von der bildungsökonomischen Forschung auf der Basis der Daten internationaler Schulleistungsstudien präsentierten Ergebnisse von Produktionsfunktionsschätzungen zur Leistungswirksamkeit von Unterschieden im *Autonomiegrad* der Schulen vermitteln ein günstige-

Schon bei der Betrachtung der deutschen Bundesländer wird deutlich, dass der Leistungsvorsprung der Schüler aus Ländern mit zentralen Prüfungen zu einem ganz erheblichen Teil auf unbeobachtete Unterschiede zwischen diesen Ländern und den Bundesländern ohne zentrale Prüfungen zurückzuführen ist.



res Bild als die dazu vorliegenden Befunde aus anderen Forschungsprogrammen. Dies gilt vor allem für die Fallstudienforschung. Zu abweichenden Ergebnissen gelangen zum Teil aber auch mit anderen Analysemethoden vorgenommene Auswertungen identischer Datensätze aus Stichprobenerhebungen. Die widersprüchlichen Befunde deuten auf eine hohe *Methodenabhängigkeit* hin. Zugleich verweisen sie auf die Notwendigkeit, den *Wirkungsmechanismen* stärkere Beachtung zu schenken, um Hinweise darauf zu erhalten, unter welchen Bedingungen eine erweiterte schulische Handlungskompetenz erfolgreich ist. Die Untersuchung von Interaktionseffekten im Rahmen des Produktionsfunktions-Ansatzes reicht dafür nicht aus. Immerhin konnte auf diese Weise aber gezeigt werden, dass die Wirksamkeit (Effektrichtung und -stärke) erweiterter schulischer

Handlungskompetenz (auch) davon abhängt, ob Schulen durch externe Leistungsprüfungen zur Rechenschaft gezogen werden oder nicht. Ausgeblendet bleibt indes die wichtige Frage, mit welchem Verhaltensrepertoire Schulen auf diese Anforderung reagieren. Durch die internationale Bildungsforschung sind *unerwünschte Verhaltensreaktionen und Handlungsfolgen* hinreichend dokumentiert. Die ökonomische Sicht der Wirksamkeit von Anreizstrukturen im Schulbereich und darauf basierende Politikempfehlungen werden dadurch problematisiert.

Auch die zu *Wettbewerbsregulativen* im Schulbereich vorliegenden Forschungsbefunde sind nicht so eindeutig, um darauf Empfehlungen für die Bildungspolitik zu gründen. Die Wirkung des Wettbewerbs durch Privatschulen ist nur im lokalen Kontext angemessen zu untersuchen. Aus dem Privatschüleranteil eines Landes lassen sich keine Rückschlüsse auf die Leistungsfähigkeit seines Schulsystems ziehen. Für die Mehrzahl der OECD-Länder – so auch für Deutschland – lässt sich bei angemessener Kontrolle des sozioökonomischen Hintergrunds der Schüler *und* der Schulen kein Leistungsvorsprung von Schülern privater Einrichtungen nachweisen.

Die plausible Annahme, dass *frühe Selektion und externe Differenzierung* von Schülern die Auseinanderentwicklung ihrer Leistungen begünstigt und das Leistungsniveau der betreffenden Schulsysteme negativ beeinflusst, lässt sich durch international vergleichende Analysen ebenfalls nicht begründen. Eine dichotome Klassifikation der verschiedenen Schulsysteme, wie sie in den besprochenen bildungsökonomischen Studien vorgenommen wird, ist der Komplexität der Schul- und Unterrichtsorganisation in den untersuchten Ländern in keiner Weise angemessen, denn: Es gibt mehrere Zeitpunkte für die Selektion von Schülern, mehrere Formen ihrer externen Differenzierung und ihrer Verteilung auf die verschiedenen Schulformen und etliche funktionale Äquivalente für hierarchisch gegliederte Schulsysteme. Schon aufgrund der problematischen Klassifikation der Systeme muss die Validität der Befunde angezweifelt werden. Wer auch nur einzelne der untersuchten Schulsysteme etwas genauer betrachtet hat oder einige Kenntnisse von der Lernkultur in den verschiedenen Ländern besitzt, der weiß im Übrigen, dass Ländervergleiche auf der Basis quantitativer Daten zu einzelnen Systemmerkmalen kein Ersatz für das Studium der Institutionen und Kulturen sind, die die Schulsysteme charakterisieren, und dass sie das Studium der dazu vorliegenden Literatur keineswegs überflüssig machen.

Literatur

Angrist & Lavy 1999: Joshua Angrist & Victor Lavy: Using Maimonides' rule to estimate the effect of class size on scholastic achievement. In: *Quarterly Journal of Economics*. Vol. 114 (2), 1999. S. 533-575.

Auswärtiges Amt 2007: Auswärtiges Amt: Türkei. Kultur- und Bildungspolitik, Medien. <http://www.auswaertiges-amt.de/diplo/de/Laenderinformationen/Tuerkei/Kultur-UndBildungspolitik.html> November 2007.

Bachmann, Iby u.a. 1996: Helmut Bachmann, Manfred Iby, Augustin Kern, Dietmar Osinger, Edwin Radnitzky, Werner Specht: Auf dem Weg zu einer besseren Schule. Evaluation der Schulautonomie in Österreich. Auswirkungen der 14. SchOG-Novelle. Wien 1996.

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

Becker 2000: Egon Becker: Von der Zukunftsinvestition zur Effektivitätskontrolle des Bildungssystems. In: Frank-Olaf Radtke & Manfred Weiß (Hrsg.): *Schulautonomie, Wohlfahrtsstaat und Chancengleichheit*. Opladen: Leske & Budrich 2000, S. 95-114.

Belfield 2001: Clive R. Belfield: International experience with demand-led financing. Education vouchers in the USA, Great Britain and Chile. New York: National Center for the Study of Privatization in Education 2001.

Belfield & Levin 2002: The effects of competition between schools on educational outcomes: A review for the United States. In: *Review of Educational Research*. Vol. 72 (2), 2002, S. 279-341.

Bifulco & Ladd 2006: Robert Bifulco & Helen F. Ladd: The impact of charter schools on student achievement: Evidence from North Carolina. In: *Education Finance and Policy*. Vol. 1 (2), 2006, S. 50-90.

Bishop & Wößmann 2004: John Bishop & Ludger Wößmann: Institutional effects in a simple model of educational production. In: *Education Economics*. Vol. 12 (1) 2004, S.17-38.

Böhlmark & Lindahl 2007: Anders Böhlmark & Mikael Lindahl: The impact of school choice on pupil achievement, segregation and costs: Swedish evidence. Bonn: IZA 2007 (Discussion Paper 2786).

Bowe, Gewirtz & Ball 1994: Richard Bowe, Sharon Gewirtz & Stephen Ball: Captured by the discourse? Issues and concerns in researching "parental choice". *British Journal of Sociology of Education*. Vol. 15 (1), S. 63-78.

Brahm 2006: Grit im Brahm: Klassengröße: eine wichtige Variable von Schule und Unterricht? <http://www.bildungsforschung.org/Archiv/2006-01/klassengroesse>

Buddin & Zimmer 2005: Richard Buddin & Ron Zimmer: Is Charter School competition in California improving the performance of traditional public schools? Santa Monica, Calif.: Rand Corporation 2005 (WR-297-EDU).

Büchel, Jürges & Schneider 2003: Felix Büchel, Hendrik Jürges & Kerstin Schneider: Die Auswirkungen zentraler Abschlussprüfungen auf die Schulleistung - Quasi-experimentelle Befunde aus der deutschen TIMSS-Stichprobe. In: *Vierteljahrshefte zur Wirtschaftsforschung*. Bd. 72 (2) 2003, S. 238-251.

Büchel, Jürges & Schneider 2004: Felix Büchel, Hendrik Jürges & Kerstin Schneider: Leistungs- und Signaleffekte zentraler Abschlussprüfungen – Eine TIMSS-Auswertung bei deutschen Haupt- und Realschülern. In: Uschi Backes-Gellner & Petra Moog: *Ökonomie der Evaluation von Schulen und Hochschulen*. Schriften des Vereins für Socialpolitik, Neue Folge. Berlin: Duncker & Humblot 2004. S. 53-73.

Bundesministerium 2002: Bundesministerium für Bildung, Wissenschaft und Kultur: Kenndaten des österreichischen Schulwesens. Ausgabe 2002. Wien: BM:BWK 2002.

Carnoy, Adamson u.a. 2007: Martin Carnoy, Frank Adamson, Amita Chudgar, Thomas F. Luschei, John F. Witte: Vouchers and public school performance. Washington: Economic Policy Institute 2007

Creemers & Reezigt 1996: Bert P. M. Creemers & Gerry J. Reezigt: School level conditions affecting the effectiveness of instruction. In: *School Effectiveness and Improvement*. Vol. 7 (3), 1996, S. 197-228.

Davies, Adnett & Mangan 2002: Peter Davies, Nick Adnett, Jean Mangan: The diversity and dynamics of competition: evidence from two local schooling markets. In: *Oxford Review of Education*. Vol. 28 (1), 2002, S. 91-107.

Davies & Quirke 2007: Scott Davies & Linda Quirke: The impact of sector on school organizations: Institutional and market logics. In: *Sociology of Education*. Vol. 80 (1), 2007, S. 66-90.

Dee 1998: Thomas S. Dee: Competition and the quality of public schools. In: *Economics of Education Review*. Vol. 17 (4), 1998, S. 419-427.

DESI-Konsortium 2007: DESI-Konsortium (Hrsg.): Sprachliche Kompetenzen. Leistungsverteilungen und Bedingungsfaktoren. Ergebnisse der DESI-Studie. Weinheim und Basel: Beltz 2007.

Eckholm 1997: Mats Eckholm: Steuerungsmodelle für Schulen in Europa: Schwedische Erfahrungen mit alternativen Ordnungsmodellen. In: *Zeitschrift für Pädagogik*. Bd. 43 (4), 1997, S. 597-608.

Fend 1998: Helmut Fend: Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lernleistung. Weinheim/München 1998.

Fend 2004: Helmut Fend: Was stimmt mit den deutschen Bildungssystemen nicht? Wege zur Erklärung von Leistungsunterschieden zwischen Bildungssystemen. In: Gundel Schümer, Klaus-Jürgen Tillmann, Manfred Weiß: Die Institution Schule und die Lebenswelt der Schüler. Wiesbaden: VS Verlag für Sozialwissenschaften 2004, S.15-38.

Fuchs & Wößmann 2007: Thomas Fuchs & Ludger Wößmann: What accounts for international differences in student performance? A re-examination using PISA data. In: Empirical Economics Vol. 32 (2) 2007, S. 433-464.

Gundlach & Wößmann 2003: Erich Gundlach & Ludger Wößmann: Bildungsressourcen, Bildungsinstitutionen und Bildungsqualität: Makroökonomische Relevanz und mikroökonomische Evidenz. Kiel: Institut für Weltwirtschaft.

Hanushek 1997: Eric A. Hanushek: Assessing the effects of school resources on student performance. An update. In: Educational Evaluation and Policy Analysis. Vol. 19 (2), 1997, S. 141-164.

Hanushek 2005: Eric. A. Hanushek: The economics of school quality. In: German Economic Review. Vol. 6 (3), 2005, S. 269-286.

Hanushek & Kimko 2000: Eric A. Hanushek & Dennis D. Kimko: Schooling, labor force quality, and the growth of nations. In: American Economic Review. Vol. 90 (5), 2000, S. 1184-1208.

Hanushek & Luke 2003: Eric A. Hanushek & Javier A. Luque: Efficiency and equity in schools around the world. In: Economics of education Review. Vol. 22 (5), 2003, S. 481-502.

Hanushek & Wößmann 2005: Eric A. Hanushek & Ludger Woessmann: Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. Bonn: Forschungsinstitut zur Zukunft der Arbeit 2005.

Hanushek & Wößmann 2007: Eric A. Hanushek & Ludger Woessmann: The role of education quality in economic growth. Washington: World Bank 2007 (World Bank Policy Research Working Paper 4122).

Heyneman 1995: Stephen P. Heyneman: Economics of education: disappointments and potential. In: Prospects. Vol. 25 (4), 1995, S. 559-593.

Holzer 2005: Thomas Holzer: PISA 2003. Einflussfaktoren auf die kantonalen Ergebnisse. Neuchâtel: Bundesamt für Statistik 2005 (Bildung und Wissenschaft 15).

Hoxby 1994: Caroline M. Hoxby: Do private schools provide competition for public schools? Cambridge, Mass. 1994 (MIT working paper).

Hoxby 2000: Caroline M. Hoxby: The effects of class size on student achievement: new evidence from population variation. In: Quarterly Journal of Economics. Vol. 115 (4), 2000. S. 1239-1285.

Husén 1967: Torsten Husén (Ed.): International study of achievement in mathematics. A comparison of twelve countries. Stockholm u.a.: Almqvist & Wiksell u.a. 1967.

Kilgore & Pendleton 1993: Sally B. Kilgore & William W. Pendleton: The organisational context of learning: Framework for understanding the acquisition of knowledge. In: *Sociology of Education*. Vol. 66 (1), 1993, S. 63-87.

Krücken 2004: Georg Krücken: Hochschulen I Wettbewerb – eine organisationstheoretische Betrachtung. In: Wolfgang Böttcher & Ewald Terhardt (Hrsg.): *Organisationstheorie in pädagogischen Feldern. Analyse und Gestaltung*. Wiesbaden: VS Verlag für Sozialwissenschaften 2004, S. 286-301.

Leithwood & Menzies 1998: Kenneth Leithwood & Teresa Menzies: Forms and effects of school-based management: a review. In: *Educational Policy*. Vol. 12 (2), 1998, S. 325-346.

Levacic & Woods 2000: Rosalind Levacic & Philip A. Woods: Quasi-markets and school performance: evidence from a study of English secondary schools. In: Manfred Weiß & Horst Weishaupt: *Bildungsökonomie und Neue Steuerung*. Frankfurt am Main u. a.: Lang 2000, S. 53-95.

Lubienski 2003: Christopher Lubienski: Innovation in education markets: Theory and evidence on the impact of competition and choice in charter schools. In: *American Educational Research Journal*. Vol. 40 (2), 2003, S. 395-443.

Lubienski & Lubienski 2006: Sarah T. Lubienski & Christopher Lubienski: School sector and academic achievement: A multilevel analysis of NAEP mathematics data. In: *American Educational Research Journal*. Vol. 43 (4), 2006, S. 651-698.

Malen, Ogawa & Kranz 1990: Betty Malen, Rodney T. Ogawa, Jennifer Kranz: What do we know about school-based management? A case study of the literature – a call for research. In: William H. Clune & John F. Witte: *choice and control in American education*. Vol. 2: The practice of choice, decentralization and school restructuring. London u. a.: Falmer Press 1990, S. 289-342.

Mansell 2007: Warwick Mansell: *Education by Numbers*. London: Politico's Publishing 2007.

Martin & Kelly 1997: Michael O. Martin & Dana L. Kelly: Third international mathematics and science study. Technical report. Vol II: Implementation and analysis – primary and middle school years. Chestnut Hill, MA: Boston College 1997.

Maslowski, Scherens & Luyten 2007: Ralf Maslowski, Jaap Scheerens, Hans Luyten: The effect of school autonomy and school internal decentralization on students' reading literacy. In: *School Effectiveness and School Improvement*. Vol. 18 (3), 2007, S. 303-334.

Murphy 2003: Joseph Murphy: Has marketization improved the quality of education? The case of charter schools.. In: *Journal of Educational Change*. Vol. 4 (1), 2003, S. 72-80.

Neal & Schanzenbach 2007: Derek Neal & Diane Whitmore Schanzenbach: Left behind by design: proficiency counts and test-based accountability. University of Chicago 2007 (paper).

OECD 2001: Organisation for Economic Co-operation and Development: Lernen für das Leben. Erste Ergebnisse der internationalen Schulleistungsstudie PISA 2000. Paris: OECD 2001.

OECD 2004: Organisation for Economic Co-operation and Development: Lernen für die Welt von Morgen: Erste Ergebnisse von PISA 2003. Paris: OECD 2004.

Picot, Dietl & Franck 1999: Arnold Picot, Helmut Dietl, Egon Franck: Organisation. Eine ökonomische Perspektive. 2. Aufl. Stuttgart: Poeschel 1999.

Robitaille 1997: David F. Robitaille (Ed.): National contexts for mathematics and science education. An encyclopedia of the education systems participating in TIMSS. Vancouver: Pacific Educational Press 1997.

Sachverständigenrat 2004: Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung: Erfolge im Ausland – Herausforderungen im Inland. Jahresgutachten 2004/05. Wiesbaden 2004.

Schneider, Carnoy u.a. 2007: Barbara Schneider, Martin Carnoy, Jeremy Kilpatrick, William H. Schmidt & Richard J. Shavelson: Estimating causal effects. American Educational Research Association 2007.

Schümer 2004: Gundel Schümer: Versuche zur Aufklärung von Leistungsunterschieden zwischen Schülern aus verschiedenen Ländern. In: Sabine Gruehn, Gerhard Kluchert & Thomas Koinzer (Hrsg.): Was Schule macht. Schule, Unterricht und Werteerziehung: theoretisch, historisch, empirisch. Weinheim & Basel: Beltz 2004. S. 113-131.

Schütz, Ursprung & Wößmann 2005: Gabriela Schütz, Heinrich W. Ursprung & Ludger Woessmann: Education policy and equality of opportunity. Bonn: Forschungsinstitut zur Zukunft der Arbeit 2005.

Schütz & Wößmann 2005: Gabriela Schütz & Ludger Wößmann: Chancengleichheit im Schulsystem: Internationale deskriptive Evidenz und mögliche Bestimmungsfaktoren. München: ifo Institut für Wirtschaftsforschung an der Universität München 2005.

Specht 1997: Werner Specht: Autonomie und Innovationsklima an Schulen: Rezeption und Wirkungen der Schulautonomie an Hauptschulen und allgemein bildenden höheren Schulen. Graz: Zentrum für Schulentwicklung 1997.

Stark Rentner, Scott u.a. 2005: Diane Stark Rentner, Caitlin Scott, Nancy Kober, Naomi Chudowsky, Victor Chudowsky, Scott Jofstus, Dalia Zabala: From the capital to the classroom: Year 4 of the No Child Left Behind Act. Washington: Center on Education Policy 2006.

Summers & Johnson 1994: Anita A. Summers & Amy W. Johnson: A review of the evidence on the effects of school-based management plans. Paper prepared for the Panel of the Economics of Education Reform and Teaching. 1994 (mimeo.)

Weiß 2002: Manfred Weiß: Stichwort Bildungsökonomie. In: Zeitschrift für Erziehungswissenschaft. Bd. 5 (2), 2002, S. 183-200.

Weiß 2007: Manfred Weiß: Bildungsfinanzierungsmodelle und ihre Umsetzung: Erfahrungen und Ergebnisse. In: Oliver Böhm-Kasper, Claudia Schuchart, Ursula Schulzeck (Hrsg.): Kontexte von Bildung. Münster: Waxmann 2007. S. 93-107.

Weiß & Timmermann 2004: Manfred Weiß & Dieter Timmermann: Bildungsökonomie und Schulstatistik. In: Werner Helsper & Jeanette Böhme (Hrsg.): Handbuch der Schulforschung. Wiesbaden. VS Verlag für Sozialwissenschaften 2004, S. 241-283.

Weiß & Preuschhoff 2006: Manfred Weiß & Corinna Preuschhoff: Gibt es einen Privatschuleffekt? In: Manfred Weiß (Hrsg.): Evidenzbasierte Bildungspolitik: Beiträge der Bildungsökonomie. Berlin. Duncker & Humblot 2006, S. 55-72.

Weiß & Bellmann 2007: Manfred Weiß & Johannes Bellmann: Bildungsfinanzierung in Deutschland und Schulqualität – eine gefährdete Balance? In: Recht der Jugend und des Bildungswesens. Bd. 55 (1), 2007, S.20-36.

West & Wößmann 2006: Martin R. West & Ludger Woessmann: Which school systems sort weaker students into smaller classes? International evidence. In: European Journal of Political Economy. Vol. 22 (4), 2006. S. 944-968.

Willms & Raudenbush 1989: Douglas J. Willms & Stephen W. Raudenbush: A longitudinal hierarchical linear model of estimating school effects and their stability. In Journal of Educational Measurement. Vol. 26, 1989, S. 209-232.

Wissenschaftsrat 2001: Wissenschaftsrat (Hrsg.): Empfehlungen zur künftigen Struktur der Lehrerbildung. Köln 2001.

Wößmann 2002: Ludger Woessmann: Central exams improve educational performance: international evidence. In: Kieler Diskussionsbeiträge. Oktober 2002.

Wößmann 2003: Ludger Woessmann: European „education production functions“: what makes a difference for student achievement in Europe? Brüssel: Europäische Kommission 2003 (Economic Papers 190).

Wößmann 2004: Ludger Woessmann: How equal are educational opportunities? Family background and student achievement in Europe and the United States. Bonn: Forschungsinstitut zur Zukunft der Arbeit 2004.

Wößmann 2005a: Ludger Woessmann: Educational production in East Asia: the impact of family background and schooling policies on student performance. In: German Economic Review. Vol. 6 (3), 2005. S. 331–353.

Wößmann 2005b: Ludger Woessmann: Educational production in Europe. In: Economic Policy. Vol. 20 (43), 2005. S. 447-504.

Wößmann 2005c: Ludger Wößmann: Kleinere Klassen = bessere Leistungen? München: Institut für Wirtschaftsforschung. ifo Schnelldienst, 58. Jg. (17), 2005. S. 6-15.

Wößmann 2005d: Ludger Wößmann: Ursachenkomplexe der PISA-Ergebnisse: Untersuchungen auf Basis internationaler Mikrodaten. In: Tertium Comparationis. Journal für International und Interkulturell Vergleichende Erziehungswissenschaft. Vol. 11 (2) 2005, S. 152-176.

Wößmann 2005e: Ludger Woessmann: The effect heterogeneity of central exams: evidence from TIMSS, TIMSS-Repeat and PISA. In: Education Economics. Vol. 13 (2) 2005, S. 143-169.

Wößmann 2005f: Ludger Wößmann: Leistungsfördernde Anreize für das Schulsystem. In: ifo-Schnelldienst Nr. 19. München: Institut für Wirtschaftsforschung 2005.

Wößmann 2007: Ludger Wößmann: Frühe Selektion führt zu mehr Chancenungleichheit. Ergebnisse nationaler und internationaler Vergleiche. In: Pädagogik, Heft 9, 2007. S. 46-51.

Wößmann & West 2006: Ludger Woessmann & Martin West: Class-size effects in school systems around the world: evidence from between-grade variation in TIMSS. In: European Economic Review. Vol. 50 (3), 2006. S. 695-736.

Ihr Kontakt zur GEW

Unsere Adressen

GEW Baden-Württemberg

Silcherstraße 7
70176 Stuttgart
Telefon: 0711/21030-0
Telefax: 0711/21030-45
www.gew-bw.de
info@gew-bw.de

GEW Bayern

Schwanthalerstraße 64
80336 München
Telefon: 089/544081-0
Telefax: 089/5389487
www.bayern.gew.de
info@gew-bayern.de

GEW Berlin

Ahornstraße 5
10787 Berlin
Telefon: 030/219993-0
Telefax: 030/219993-50
www.gew-berlin.de
info@gew-berlin.de

GEW Brandenburg

Alleestraße 6a
14469 Potsdam
Telefon: 0331/27184-0
Telefax: 0331/27184-30
www.gew-brandenburg.de
info@gew-brandenburg.de

GEW Bremen

Löningstraße 35
28195 Bremen
Telefon: 0421/33764-0
Telefax: 0421/33764-30
www.gew-bremen.de
info@gew-hb.de

GEW Hamburg

Rothenbaumchaussee 15
20148 Hamburg
Telefon: 040/414633-0
Telefax: 040/440877
www.gew-hamburg.de
info@gew-hamburg.de

GEW Hessen

Zimmerweg 12
60325 Frankfurt am Main
Telefon: 069/971293-0
Telefax: 069/971293-93
www.gew-hessen.de
info@gew-hessen.gew.de

GEW Mecklenburg-Vorpommern

Lübecker Straße 265a
19059 Schwerin
Telefon: 0385/485270
Telefax: 0385/4852724
www.gew-mv.de
Landesverband@mvp.gew.de

GEW Niedersachsen

Berliner Allee 16
30175 Hannover
Telefon: 0511/33804-0
Telefax: 0511/33804-46
www.gew-nds.de
email@gew-nds.de

GEW Nordrhein-Westfalen

Nünningstraße 11
45141 Essen
Telefon: 0201/294030-1
Telefax: 0201/29403-51
www.gew-nrw.de
info@gew-nrw.de

GEW Rheinland-Pfalz

Neubrunnenstraße 8
55116 Mainz
Telefon: 06131/28988-0
Telefax: 06131/28988-80
www.gew-rlp.de
gew@gew-rlp.de

GEW Saarland

Mainzer Straße 84
66121 Saarbrücken
Telefon: 0681/66830-0
Telefax: 0681/66830-17
www.gew-saarland.de
info@gew-saarland.de

GEW Sachsen

Nonnenstraße 58
04229 Leipzig
Telefon: 0341/4947404
Telefax: 0341/4947406
www.gew-sachsen.de
gew-sachsen@t-online.de

GEW Sachsen-Anhalt

Markgrafenstraße 6
39114 Magdeburg
Telefon: 0391/73554-0
Telefax: 0391/7313405
www.gew-lsa.de
lv@gew-lsa.de

GEW Schleswig-Holstein

Legienstraße 22-24
24103 Kiel
Telefon: 0431/554220
Telefax: 0431/554948
www.gew-sh.de
info@gew-sh.de

GEW Thüringen

Heinrich-Mann-Straße 22
99096 Erfurt
Telefon: 0361/59095-0
Telefax: 0361/59095-60
www.gew-thueringen.de
info@gew-thueringen.de

GEW-Hauptvorstand

Reifenberger Straße 21
60489 Frankfurt am Main
Telefon: 069/78973-0
Telefax: 069/78973-201
www.gew.de
info@gew.de

GEW-Hauptvorstand Parlamentarisches Verbindungsbüro Berlin

Wallstraße 65
10179 Berlin
Telefon: 030/235014-11 bis 15
Telefax: 030/235014-10
info@buero-berlin.gew.de

Die GEW im Internet

www.gew.de

Antrag auf Mitgliedschaft

(Bitte in Druckschrift ausfüllen)

Persönliches

Frau/Herr

Nachname (Titel)

Vorname

Straße, Nr.

Postleitzahl, Ort

Telefon / E-Mail

Geburtsdatum

Nationalität

gewünschtes Eintrittsdatum

bisher gewerkschaftlich organisiert bei von/bis (Monat/Jahr)

Name/Ort der Bank

Kontonummer

BLZ

Berufliches

Berufsbezeichnung für Studierende: Berufsziel

Fachgruppe

Diensteintritt / Berufsbeginn

Tarif- / Besoldungsgebiet

Tarif / Besoldungsgruppe

Stufe

seit

Bruttoeinkommen € monatlich (falls nicht öffentlicher Dienst)

Betrieb / Dienststelle

Träger des Betriebs / der Dienststelle

Straße, Nr. des Betriebs / der Dienststelle

Postleitzahl, Ort des Betriebs/der Dienststelle

Ihr Mitgliedsbeitrag:

- Beamtinnen und Beamte zahlen 0,75 Prozent der 6. Stufe.
- Angestellte zahlen 0,7 Prozent der Entgeltgruppe und Stufe, nach der vergütet wird.
- Der Mindestbeitrag beträgt immer 0,6 Prozent der untersten Stufe der Entgeltgruppe I des TVöD.
- Arbeitslose zahlen ein Drittel des Mindestbeitrages.
- Studierende zahlen einen Festbetrag von 2,50 Euro.
- Mitglieder im Referendariat oder Praktikum zahlen einen Festbetrag von 4 Euro.
- Mitglieder im Ruhestand zahlen 0,66 Prozent ihrer Ruhestandsbezüge.

Weitere Informationen sind der Beitragsordnung zu entnehmen.

Jedes Mitglied der GEW ist verpflichtet, den satzungsgemäßen Beitrag zu entrichten. Der Austritt ist mit einer Frist von drei Monaten schriftlich dem Landesverband zu erklären und nur zum Ende eines Kalendervierteljahres möglich.

Mit meiner Unterschrift auf diesem Antrag ermächtige ich die GEW zugleich widerruflich, den von mir zu leistenden Mitgliedsbeitrag vierteljährlich von meinem Konto abzubuchen.

Beschäftigungsverhältnis:

- | | |
|---|--|
| <input type="checkbox"/> Honorarkraft | <input type="checkbox"/> in Elternzeit |
| <input type="checkbox"/> angestellt | <input type="checkbox"/> befristet bis _____ |
| <input type="checkbox"/> beurlaubt ohne Bezüge | <input type="checkbox"/> teilzeitbeschäftigt mit ____ Std. / Woche |
| <input type="checkbox"/> beamtet | <input type="checkbox"/> teilzeitbeschäftigt mit ____ Prozent |
| <input type="checkbox"/> in Rente / pensioniert | <input type="checkbox"/> Referendariat / Berufspraktikum |
| <input type="checkbox"/> im Studium | <input type="checkbox"/> arbeitslos |
| <input type="checkbox"/> Altersteilzeit | <input type="checkbox"/> Sonstiges _____ |

Mit meiner Unterschrift auf diesem Antrag ermächtige ich die GEW zugleich widerruflich, den von mir zu leistenden Mitgliedsbeitrag vierteljährlich von meinem Konto abzubuchen.

Ort, Datum

Unterschrift

wird von der GEW ausgefüllt

GEW-KV/-OV

Dienststelle

Fachgruppe

Kassiererstelle

Tarfbereich

Beschäftigungsverhältnis

Mitgliedsbeitrag €

Startmonat

Bitte senden Sie den ausgefüllten Antrag an den für Sie zuständigen Landesverband der GEW bzw. an den Hauptvorstand.

Die uns von Ihnen angegebenen personenbezogenen Daten sind nur zur Erfüllung unserer satzungsgemäßen Aufgaben auf Datenträgern gespeichert und entsprechend den Bestimmungen des Bundesdatenschutzgesetzes geschützt.

Vielen Dank!
Ihre GEW

