

## Zentrale Prüfungen als „Währung“ des Bildungssystems: Zur Komplementarität von Schulautonomie und Zentralprüfungen

Von Ludger Wößmann\*

**Zusammenfassung:** So wie die Währung im Wirtschaftssystem die Funktion eines „Wertmaßstabes“ innehat, können einheitliche Leistungsüberprüfungen im Schulsystem als Bewertungsmaßstab fungieren, der zur Überwindung von Informationsasymmetrien beiträgt und opportunistisches Verhalten dezentraler Entscheidungsträger verhindert. Damit werden Zentralprüfungen zur Voraussetzung dafür, dass dezentral organisierte Schulsysteme zu hohen Schülerleistungen führen können. Diese Komplementarität zwischen Zentralprüfungen und Schulautonomie wird zunächst in einem Prinzipal-Agenten-Ansatz der Bildungsproduktion theoretisch abgebildet und dann anhand der internationalen TIMSS-Schülerleistungsvergleichsstudien empirisch getestet. Die mikroökonomischen Schätzungen weisen in einem ersten Schritt starke positive Leistungseffekte zentraler Prüfungen nach. In einem zweiten Schritt werden Interaktionseffekte zwischen Zentralprüfungen und Schulautonomie zugelassen. Dabei zeigt sich, dass sich Schulautonomie in Schulsystemen ohne Zentralprüfungen vielfach negativ auf Schülerleistungen auswirkt. In Schulsystemen mit Zentralprüfungen werden diese negativen Autonomieeffekte zumeist vollständig abgebaut, und im Fall der Schulautonomie über Lehrgelöhner kehren sie sich in stark positive Effekte um. Eine effiziente Bildungspolitik sollte also Zentralprüfungen mit Schulautonomie verbinden, d. h. Standards extern vorgeben und überprüfen, aber den Schulen die Art ihres Erreichens überlassen.

**Summary:** Just as currencies serve as unit of value in economic systems, central exams can serve as measures of value in school systems, thereby mitigating informational asymmetries and preventing opportunistic behavior in decentralized decision-making. Central exams are thus a prerequisite for decentralized school systems to achieve high student performance. I first sketch out this complementarity between central exams and school autonomy in a principal-agent model of educational production and then test it empirically using the TIMSS international student achievement tests. In a first step, the microeconomic estimations reveal large positive effects of central exams on student performance. In a second step, I allow for interaction effects between central exams and school autonomy. In school systems without central exams, school autonomy often exerts a negative impact on student performance. Central exams remove these negative effects of autonomy, and turn them around into positive effects in the case of school autonomy in salary decisions. Efficient education policies would thus combine central exams with school autonomy, externally setting and testing standards but leaving it to schools how to pursue them.

\* CESifo München, ifo Institut für Wirtschaftsforschung, Poschingerstr. 5, 81679 München, E-Mail: woessmann@ifo.de, Internet: [www.cesifo.de/link/woessmann\\_1.htm](http://www.cesifo.de/link/woessmann_1.htm)

## 1 Einleitung

Eine hohe Qualität der in Schulen vermittelten Bildung bewirkt eine höhere Produktivität und eine ausgeglichene Einkommensverteilung von Volkswirtschaften (Wößmann 2003a, Gundlach et al. 2003). Deshalb ist es eine (nicht nur) aus volkswirtschaftlicher Sicht entscheidende Frage, wie sich eine bessere Qualität der schulischen Bildung sicherstellen lässt. Zahlreiche empirische Evidenz legt nahe, dass dies nicht einfach durch eine Erhöhung der Bildungsausgaben zu erreichen ist: In den meisten Fällen scheinen zusätzliche Ressourcen weder über die Zeit noch im Querschnitt zu verbesserten Schülerleistungen zu führen (vgl. Gundlach et al. 2001, Hanushek 2002 und 2003, Wößmann und West 2002). Stattdessen geht aber eine institutionelle Gestaltung des Schulsystems, die mehr leistungsfördernde Anreize setzt, mit besseren Schülerleistungen einher (Wößmann 2003b).

Eine solche als leistungsfördernd identifizierte Institution sind zentrale bzw. externe Prüfungen (Bishop 1997, Wößmann 2003b). Anstatt die Gestaltung, Ausführung und Benotung der schulischen Leistungsüberprüfung den einzelnen Lehrern oder Schulen zu überlassen, werden Zentralprüfungen von einer externen Instanz durchgeführt. Die erzeugte, allgemein vergleichbare Leistungsinformation verändert den Informationsstand im Schulsystem. Durch diese Informationsbereitstellung verändern Zentralprüfungen die Anreize der am Bildungsprozess beteiligten Personengruppen: Es ist nun wahrnehmbar, welche Leistungen Schüler, Lehrer und Schulen erreicht haben, so dass die Basis für entsprechende Konsequenzen gegeben ist. Dies erzeugt leistungsfördernde Anreize für alle Beteiligten (vgl. Bishop und Wößmann 2003). Dafür kommt es nicht darauf an, ob die „zentralen“ Prüfungen von einer nationalen Behörde, von Behörden auf Bundesländerebene oder einheitlich von privaten Anbietern durchgeführt werden; entscheidend ist, dass sie „extern“ in Bezug auf die einzelne Schule stattfinden.

Der vorliegende Beitrag untersucht, welche Auswirkungen diese Anreizwirkung von Zentralprüfungen auf die relative Leistungsfähigkeit dezentral organisierter Schulsysteme hat. Die zentrale Botschaft des Beitrags ist, dass Zentralprüfungen – trotz der scheinbaren Konnotation eines zentral organisierten Systems – keineswegs notwendigerweise Bestandteil eines zentral regulierten Schulsystems sein müssen, sondern dass sie im Gegenteil geradezu eine Voraussetzung dafür sind, dass ansonsten dezentral organisierte Schulsysteme effizient funktionieren können. Dies hat damit zu tun, dass im Bildungssystem aufgrund asymmetrischer Informationsverteilungen und unterschiedlicher Interessenlagen zwischen Auftraggebern und Handelnden vielfach starke Anreize für opportunistisches Verhalten vorliegen. Solange lokale Entscheidungsträger in solchen Fällen für ihr Verhalten nicht zur Rechenschaft gezogen werden können, da keine Information über die erzielten Leistungen vorliegt, wird sich eine dezentrale Entscheidungsautonomie der Schulen negativ auf die erzielten Schülerleistungen auswirken. Sobald aber Zentralprüfungen die asymmetrischen Informationen überwinden und opportunistisches Verhalten offen legen, so dass dezentral handelnde Schulen für ihre erzielten Leistungen gerade stehen müssen, können sich die negativen Autonomieeffekte aufgrund von nun zum Tragen kommenden dezentralen Wissensvorsprüngen in positive Effekte umkehren.

In diesem Sinne besteht also eine Komplementarität zwischen Schulautonomie und Zentralprüfungen: Erst aufgrund der durch Zentralprüfungen bedingten leistungsfördernden Anreizgestaltung führt Schulautonomie zu besseren Bildungsleistungen, und aufgrund der durch Schulautonomie möglichen Ausnutzung dezentraler Wissensvorsprünge können

Zentralprüfungen zur Erzeugung von besonders guten Bildungsleistungen beitragen. Dies bedeutet auch, dass die vielfach geforderte Dezentralisierung des Schulsystems (vgl. etwa Weiß 1998, World Bank 1999: 49–50) nur dann zu positiven Leistungseffekten führen wird, wenn etwa durch zentrale Prüfungen sichergestellt ist, dass die dezentralen Entscheidungsträger Anreize zu leistungsförderndem Verhalten haben.<sup>1</sup>

Insofern können Zentralprüfungen im Schulsystem eine Funktion ausüben, die (ebenfalls zentral bereitgestellte) Währungen im Wirtschaftssystem ausüben: So wie Geld im Wirtschaftssystem die Funktion einer „Recheneinheit“ innehat (vgl. z. B. Issing 1998), können einheitliche Leistungsüberprüfungen im Schulsystem eine „Recheneinheitsfunktion“ übernehmen. Durch seine Funktion als Recheneinheit oder „Wertmaß“ ermöglicht es das Geld, die Schätzung des Gebrauchswertes eines Gegenstandes zu präzisieren und zu quantifizieren und ihn somit mit allen Alternativen vergleichbar zu machen (vgl. etwa Schumpeter 1970: 25–35). Solche quantitativen Leistungs- und Bewertungsmaßstäbe sind „offenbar für die Rationalisierung des Handelns ... von der allergrößten Bedeutung, von ähnlicher Bedeutung ... wie Sprache und Schrift“ (Schumpeter 1970: 27). In ähnlicher Weise können Zentralprüfungen eine entscheidende Rolle als Leistungs- und Bewertungsmaßstab im Schulsystem spielen.

Darüber hinaus haben Brunner und Meltzer (1971) in ihrer Theorie des Geldes die besondere Rolle betont, die der Währung dadurch zukommt, dass sie zur Überwindung asymmetrischer Informationen in einer Tauschwirtschaft beiträgt. In Tauschwirtschaften können Informationen über Marktpreise und Güterqualitäten nicht kostenlos gewonnen werden. Geld verringert die Kosten der Informationsgewinnung, und „It is the uneven distribution of information ... that induces individuals to search for, and social groups to accept, alternatives to barter“ – nämlich Geld (Brunner und Meltzer 1971: 786). So wie eine entscheidende Funktion des Geldes also in der Reduktion von Transaktionskosten bei asymmetrischer Informationsverteilung besteht, so können auch Zentralprüfungen als standardisierte Maßeinheit im Schulsystem zur Überwindung asymmetrischer Informationen zwischen Angebots- und Nachfrageseite beitragen. Wie das zentral bereitgestellte Geld im Wirtschaftssystem werden damit zentrale Prüfungen im Schulsystem zur Voraussetzung für die Funktionsfähigkeit eines dezentral organisierten Systems von autonom handelnden Agenten.<sup>2</sup>

Diese Funktion von Zentralprüfungen als „Währungseinheit des Bildungssystems“ soll im vorliegenden Beitrag zunächst theoretisch und dann empirisch untersucht werden. Die theoretische Analyse bildet Zentralprüfungen als Monitoring-Mittel in einer Prinzipal-Agenten-Theorie des Bildungssystems ab. Der Prinzipal-Agenten-Modellrahmen beachtet neben den Vorteilen lokalen Wissens auch die Gefahren lokalen opportunistischen Verhaltens. Damit hängen die Auswirkungen lokaler Autonomie davon ab, wie stark die jeweiligen Möglichkeiten für opportunistisches Verhalten sind. In Entscheidungsbereichen wie Budget- und Gehaltsfragen, in denen aufgrund von divergierenden Interessenlagen starke Anreize für opportunistisches Verhalten vorliegen, kommt es entscheidend darauf an, ob Zentral-

<sup>1</sup> In diesem Lichte ist es auch wenig überraschend, dass Summers und Johnson (1996) in einem Überblick über Dezentralisierungsreformen in den Vereinigten Staaten finden, dass Dezentralisierung sich nicht immer positiv auswirkt.

<sup>2</sup> Allerdings sollten die Parallelen zwischen Zentralprüfungen im Schulsystem und Währungen im Wirtschaftssystem auch nicht zu weit getrieben werden: Währungen führen im Wirtschaftssystem noch weitere Funktionen aus, indem sie etwa als Tausch- und Zahlungsmittel und als Wertaufbewahrungsmittel fungieren, und das von Zentralprüfungen wiedergegebene Wertmaß gibt nur in bedingtem Maße einen *Tauschwert* an.

prüfungen vorhanden sind, die es aufgrund der von ihnen bereitgestellten Informationen ermöglichen, dass die lokalen Entscheidungsträger für ihr Verhalten zur Rechenschaft gezogen werden können. Je nachdem, ob Zentralprüfungen im System vorhanden sind oder nicht, werden die positiven Leistungseffekte von Schulautonomie aufgrund der Verwertung dezentralen Wissens ihre negativen Effekte aufgrund von dezentralem opportunistischen Verhalten übersteigen oder nicht.

Die anschließende empirische Analyse weist diese Komplementarität von Schulautonomie und Zentralprüfungen anhand des internationalen Mikrodatensatzes der Schülerleistungstests TIMSS und TIMSS-Repeat, der nahezu eine halbe Million Mittelstufenschüler aus 54 Ländern umfasst, empirisch nach. Dazu werden in die Schätzung internationaler Bildungsproduktionsfunktionen Interaktionseffekte zwischen Schulautonomie und Zentralprüfungen eingeführt. Es zeigt sich etwa, dass Schulautonomie in Gehaltsfragen in Schulsystemen ohne Zentralprüfungen einen statistisch signifikanten negativen Einfluss auf die erzielten Schülerleistungen ausübt, der sich in Systemen mit Zentralprüfungen in einen statistisch signifikanten positiven Einfluss umkehrt. So weist die internationale Evidenz in verschiedenen Entscheidungsbereichen eine Komplementarität zwischen zentralen Prüfungen und dezentraler Entscheidungsfindung nach. Zumeist schneiden solche Schulsysteme am besten ab, die zentrale Prüfungen mit Schulautonomie verbinden, die also Standards vorgeben und ihr Erreichen überprüfen, aber es gleichzeitig den Schulen überlassen, *wie* die extern gesetzten Standards zu erreichen sind.

## 2 Opportunismus, lokales Wissen, Zentralprüfungen und Schulautonomie

### 2.1 Ein Prinzipal-Agenten-Ansatz der Bildungsproduktion

In theoretischer Betrachtungsweise lässt sich die Bildungsproduktion als ein Netzwerk von Prinzipal-Agenten-Beziehungen verstehen, in denen ein Prinzipal (z. B. Eltern) einen Agenten (z. B. einen Schulleiter) damit beauftragt, eine Leistung (die Bildung des Kindes) für den Prinzipal durchzuführen. Laffont und Martimort (2002: 2) beschreiben dezentrale Information und eine widerstreitende Interessenkonstellation als die beiden wesentlichen Bestandteile der Anreizwirkungen, die solche Prinzipal-Agenten-Beziehungen zum Problem machen: „Delegation of a task to an agent who has different objectives than the principal who delegates this task is problematic when information about the agent is imperfect.“ Hat nämlich der Agent Interessen, die von denen des Prinzipals abweichen, und bestehen asymmetrische Informationen darüber, was der Agent tatsächlich leistet, so kann der Agent seine eigenen Interessen anstatt der Interessen des Prinzipals verfolgen, ohne dass der Prinzipal dies feststellen und sanktionieren kann.

Zentrale Prüfungen können dem Problem von unvollständiger Überprüfung („Monitoring“) der Handlungen der Agenten im Bildungssystem ein Stück weit beikommen, indem sie Informationen über die Leistungen der einzelnen Schüler relativ zur nationalen (oder regionalen) Schülerpopulation erzeugen. Indem sie die der Prinzipal-Agenten-Beziehung inhärenten Monitoring-Probleme verringern, bringen sie die Anreize der Agenten stärker mit den Interessen des Prinzipals und damit mit den Zielen des Bildungssystems in Einklang (vgl. Wößmann 2002b). Sie machen den Leistungsstand der Schüler für Eltern, Lehrer, potentielle Arbeitgeber und weiterführende Bildungseinrichtungen sichtbar und vergleichbar, so dass bessere Leistungen belohnt werden können; sie verhindern, dass ganze Wis-

sensgebiete in einzelnen Klassen ohne Konsequenzen für die Benotung ausgelassen werden können; und sie eröffnen Eltern und Schulleitern, ob die Lehrer eine erfolgreiche Wissensvermittlung leisten.

## 2.2 Schulautonomie mit und ohne Monitoring

Schulautonomie oder die Dezentralisierung von Entscheidungsmacht kann als eine solche Delegation einer Aufgabe von einem Prinzipal, der die Wissenserzeugung im Schulsystem erreichen will, auf Agenten, nämlich die Schulen, verstanden werden. Nicht immer muss dies ein „Problem“ sein: Solange es keine unterschiedlichen Interessen oder keine asymmetrische Information gibt, kann ein zielkonformes Verhalten der Agenten sichergestellt werden. Nur wenn beides gegeben ist, bestehen für die Agenten Anreize und Möglichkeiten zu opportunistischem Verhalten, ohne dass die Gefahr bestünde, dass solches Verhalten festgestellt und sanktioniert würde.

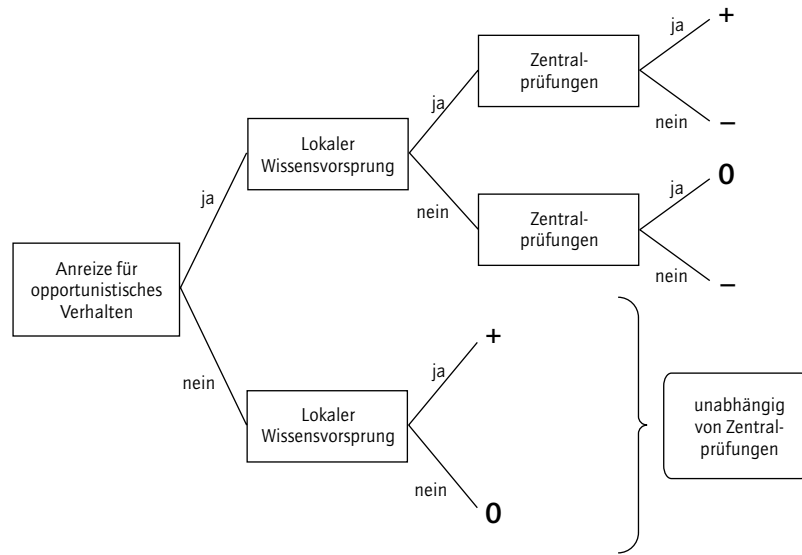
Die Gefahr des Opportunismus bei dezentraler Entscheidungsmacht ist also auf Entscheidungsbereiche beschränkt, in denen die Interessenlage der dezentralen Entscheidungsträger von einer Steigerung des Wissens der Schüler abweicht. Dies ist z. B. immer dann vorstellbar, wenn es bei der Entscheidung um die finanzielle Stellung oder die zu erfüllende Arbeitslast der Schulen geht: In solchen Fällen ist es für die schulischen Entscheidungsträger rational, ihr eigenes Interesse anstelle der Leistungsförderung der Schüler zu verfolgen, solange mögliche überwachende Instanzen wie die Schulaufsicht oder die Eltern keine Informationen über das konkrete Verhalten der Schulen haben. Aufgrund des dezentralen Produktionscharakters der Bildungsproduktion ist aber nahezu immer ein hohes Maß an Informationsasymmetrie über das schulische Verhalten gegeben. Durch die Bereitstellung von Informationen über die erzeugten Leistungen können Zentralprüfungen diese Informationsasymmetrie aber zumindest teilweise überwinden. So kommt Zentralprüfungen eine erhebliche Effizienzwirkung in der Bildungsproduktion zu, wann immer divergierende Interessenlagen in einer Entscheidungssituation opportunistisches Verhalten wahrscheinlich machen.

Bei der Betrachtung von Schulautonomie im Bildungssystem kommt noch ein zweiter wichtiger Punkt hinzu: In vielen Entscheidungsbereichen wissen dezentrale Entscheidungsträger viel besser, als es eine zentrale Instanz jemals könnte, wie die Bildungsproduktion am effizientesten gestaltet werden kann. So haben Lehrer zumeist einen lokalen Wissensvorsprung, wenn es darum geht, wie ein bestimmtes Wissensgebiet ihren spezifischen Schülern am besten nahe gebracht werden kann. Dies ist nur ein Beispiel des weit verbreiteten „knowledge of the particular circumstances of time and place“ (Hayek 1945: 522), das eine dezentrale Leistungsbereitstellung weit effizienter als eine zentrale Planung machen kann. Entscheidend ist aber nun, ob die dezentralen Entscheidungsträger auch Anreize verspüren, diesen lokalen Wissensvorsprung in der Bildungsproduktion nutzbar zu machen. Dies wird nur dann der Fall sein, wenn es anderen überhaupt auffällt, ob sich die dezentralen Entscheidungsträger die Mühe gemacht haben, ihr lokales Wissen zu nutzen – also nur dann, wenn Informationsasymmetrien etwa durch zentrale Prüfungen überbrückt werden.

Abbildung 1 stellt die entsprechenden Leistungswirkungen von Schulautonomie für verschiedene Entscheidungsbereiche dar, die sich durch das Bestehen oder Nichtbestehen von Anreizen für opportunistisches Verhalten und durch das Bestehen oder Nichtbestehen von

Abbildung 1

### Auswirkungen von Autonomie auf Schülerleistungen



„Anreize für opportunistisches Verhalten“ und „lokaler Wissensvorsprung“ sind Merkmale des jeweiligen Entscheidungsbereichs, der autonom oder nicht autonom organisiert werden kann.

- + Autonomie ist leistungssteigernd.
- Autonomie ist leistungsmindernd.
- 0 Kein Leistungsunterschied zwischen autonomer und zentraler Entscheidung.

Quelle: Eigene Darstellung.

lokalen Wissensvorsprüngen charakterisieren lassen. Für den Fall von Entscheidungsbereichen, in denen keine Anreize für opportunistisches Verhalten bestehen, da es keine Interessendivergenz zwischen Agenten und Prinzipal gibt, lassen sich die Leistungswirkungen von Schulautonomie sehr einfach ableiten: Wenn lokale Entscheidungsträger in solchen Entscheidungsbereichen einen Wissensvorsprung haben, so wirkt sich Schulautonomie positiv auf die erzeugten Bildungsleistungen aus, da zwar die Vorteile lokaler Entscheidungsfindung (lokaler Wissensvorsprung) bestehen, nicht aber die Nachteile (opportunistisches Verhalten). Wenn lokale Entscheidungsträger in den Entscheidungsbereichen keinen Wissensvorsprung haben, dann gibt es folglich keinen Unterschied zwischen dezentral und zentral organisierter Entscheidungsfindung. In beiden Fällen ist es für den Einfluss der Schulautonomie auf die Schülerleistungen unerheblich, ob es in dem Schulsystem zentrale Prüfungen gibt oder nicht: Es besteht ja definitionsgemäß keine Gefahr für opportunistisches Verhalten, die gebannt werden müsste.

Die Existenz von Zentralprüfungen ist nur in solchen Entscheidungsbereichen für die Leistungswirkung von Schulautonomie von Bedeutung, in denen Anreize für opportunistisches Verhalten bestehen, da die Interessen der Agenten von denen des Prinzipals abweichen. Betrachten wir zunächst den Fall, dass es in solchen Entscheidungsbereichen keinen lokalen Wissensvorsprung und deshalb keine Vorteile dezentraler Entscheidungsfindung gibt. Gibt es in dem Schulsystem keine Zentralprüfungen, so wirkt sich Schulautonomie in diesen Entscheidungsbereichen negativ auf die erzielten Schülerleistungen aus, weil die de-

zentrale Entscheidung – im Gegensatz zur zentralen Entscheidung – zu opportunistischem Handeln führt. Gibt es demgegenüber aber Zentralprüfungen, so sind die Gefahren lokalen opportunistischen Handelns und damit die negativen Leistungseffekte auch bei dezentraler Entscheidungsfindung gebannt, so dass es keine Leistungsunterschiede zwischen Schulautonomie und zentraler Entscheidung gibt.

Bestehen in dem betrachteten Entscheidungsbereich schließlich sowohl Anreize für opportunistisches Verhalten als auch lokale Wissensvorteile, so kommt es für die Leistungswirkungen von dezentralen Entscheidungen wiederum auf die Existenz von Zentralprüfungen an. Sind Zentralprüfungen gegeben, so sind die Nachteile opportunistischen Verhaltens gebannt, so dass die lokalen Wissensvorteile insgesamt zu einem positiven Leistungseffekt der Schulautonomie führen dürften. Sind Zentralprüfungen aber nicht gegeben, so stehen den lokalen Wissensvorteilen die Nachteile opportunistischen Verhaltens gegenüber, und für den Gesamteffekt der Schulautonomie kommt es auf die relative Größe dieser beiden Teileffekte an. A priori ist es also nicht eindeutig, ob sich in diesen Entscheidungsbereichen ein geringer positiver, gar kein oder ein negativer Gesamteffekt der Schulautonomie ergibt; aufgrund der im nächsten Abschnitt berichteten empirischen Ergebnisse, in denen der negative Opportunismuseffekt den positiven Wissensseffekt zumeist zu übersteigen scheint, ist in Abbildung 1 ein negativer Gesamteffekt angegeben. In diesem Fall drehen zentrale Prüfungen einen ursprünglich negativen Leistungseffekt von Schulautonomie gänzlich in einen positiven Leistungseffekt um.

### 3 Internationale Evidenz

#### 3.1 Die TIMSS-Datensätze

Um die theoretisch abgeleiteten Hypothesen empirisch testen zu können, benutze ich die Daten der internationalen Schülerleistungsvergleichsstudien „Third International Mathematics and Science Study“ (TIMSS). Die TIMS-Studie wurde zunächst 1995 durchgeführt („TIMSS-95“), und 1999 gab es eine Wiederholungsstudie („TIMSS-Repeat“). Während für TIMSS-95 international vergleichbare Daten für 266 545 Schüler aus 6 107 Schulen in 39 Ländern vorliegen, sind es für TIMSS-Repeat 180 544 Schüler aus 6 068 Schulen in 38 Ländern. Der zusammengesetzte („gepoolte“) Datensatz enthält also insgesamt 447 089 Schüler- und 77 Länderobservationen, und da nur 23 Länder an beiden Tests teilgenommen haben, beinhaltet der gepoolte Datensatz 54 verschiedene Länder.<sup>3</sup>

Beide TIMS-Studien wurden von der International Association for the Evaluation of Educational Achievement (IEA) durchgeführt. In der TIMSS-95-Mittelstufenstudie wurden die beiden Jahrgangsstufen mit dem größten Anteil 13-Jähriger getestet, was in den meisten Ländern der siebten und achten Jahrgangsstufe entspricht. In TIMSS-Repeat wurde die obere dieser beiden Jahrgangsstufen getestet. Innerhalb jedes Landes wurde eine repräsentative Stichprobe von rund 150 Schulen erhoben, innerhalb derer jeweils eine (zufällig ausgewählte) achte Klasse und – in TIMSS-95 – eine siebte Klasse komplett getestet wurden.

<sup>3</sup> Für nähere Informationen zu den beiden TIMS-Studien siehe etwa Gonzalez und Smith (1997) und Gonzalez und Miles (2001). Wößmann (2003b, 2002b) enthält weiterführende Informationen und Verweise zu den speziellen, im vorliegenden Beitrag verwendeten Datensätzen.

In der vorliegenden Untersuchung nutze ich die individuellen Schülerdaten des gepoolten Datensatzes, so dass ich sowohl möglichst viele verschiedene Schulsysteme mit und ohne Zentralprüfungen als auch lokale Unterschiede in dem Ausmaß der Schulautonomie innerhalb dieser Schulsysteme betrachten kann. Neben den Daten über die Mathematik- und Naturwissenschaftsleistungen der einzelnen Schüler beinhalten die TIMSS-Datensätze umfangreiche Hintergrundinformationen, die in verschiedenen Hintergrundfragebögen erhoben wurden. So erlauben es Daten aus Schülerfragebögen, in den Schätzungen für umfangreiche Einflüsse des persönlichen und familiären Hintergrunds der Schüler zu kontrollieren. Lehrerfragebögen enthalten Daten sowohl über Lehrermerkmale und die Ressourcenausstattung der Klassen als auch über die Einflussmöglichkeiten der Lehrer in verschiedenen Entscheidungsbereichen. Schulleiterfragebögen geben schließlich vor allem Auskunft über das Ausmaß von Schulautonomie in verschiedenen Entscheidungsbereichen.

Zusätzlich zu diesen TIMSS-Daten enthält der hier verwendete Datensatz Informationen darüber, ob in den betreffenden Ländern (oder in Regionen innerhalb der Länder) zentrale Abschlussprüfungen am Ende der Sekundarstufe stattfinden. Dabei werden alle Formen so genannter „curriculum-based external exit exam systems“ (Bishop 1997) berücksichtigt, nicht aber universitäre Eingangstests, da zumeist nicht alle Schüler an diesen teilnehmen und sie damit keinen integralen Bestandteil des Schulsystems darstellen. Die verwendeten Informationen über Zentralprüfungen stammen aus vergleichenden Bildungsstudien, Bildungsenzyklopädien, Interviews mit Repräsentanten der nationalen Schulsysteme, Regierungsdokumenten und Hintergrundpapieren. In Fällen, in denen Zentralprüfungen nur in einigen Regionen eines Landes stattfinden, geben die verwendeten Daten den Anteil der Schüler an, die an den Zentralprüfungen teilnehmen.

### 3.2 Zentrale Prüfungen und Schülerleistungen

Bevor im folgenden Abschnitt der Unterschied von Autonomieeffekten in Schulsystemen mit und ohne Zentralprüfungen untersucht wird, soll zunächst der allgemeine Effekt von Zentralprüfungen auf die erzielten Schülerleistungen geschätzt werden, indem die Leistungen von Schülern in Schulsystemen mit Zentralprüfungen mit den Leistungen von Schülern in Schulsystemen ohne Zentralprüfungen international verglichen werden. Der Effekt  $\alpha$  von Zentralprüfungen  $Z$  wird mit Hilfe folgender Spezifikation geschätzt:

$$T_{iksl} = \alpha Z_l + A_{ksl} \beta + C_{iksl} \gamma + \mu_l + \eta_s + \nu_k + \varepsilon_i, \quad (1)$$

wobei  $Z_l$  der Anteil der Schüler in Land  $l$  ist, der an Zentralprüfungen teilnimmt.<sup>4</sup>  $T_{iksl}$  ist das Testergebnis des Schülers  $i$  in Klasse  $k$  in Schule  $s$  in Land  $l$ . Die TIMSS-Testergebnisse sind so skaliert, dass sie in jedem Fach einen internationalen Mittelwert von 500 und eine internationale Standardabweichung von 100 aufweisen. Neben zwölf Indikatoren der Schulautonomie  $A$  in verschiedenen Entscheidungsbereichen kontrolliert die Schätzung auch für einen umfangreichen Kontrollvariablenvektor  $C$ , der 17 Variablen für den persönlichen und familiären Hintergrund der Schüler, 13 Variablen für die schulische Ressourcenausstattung und Lehrermerkmale und sechs Variablen für weitere institutionelle Merkmale der Schulsysteme wie Zentralität von Curriculum und Lehrbuchfestsetzung um-

<sup>4</sup> Da es sich meistens um nationale Zentralprüfungssysteme handelt, nimmt  $Z$  zumeist die Dummywerte 0 oder 1 an.



Tabelle 1

**Der Effekt von Zentralprüfungen auf Schülerleistungen**

	Mathematik	Naturwissenschaften	Schüler	Länder
TIMSS-95	40,9* (13,5)	39,7* (9,9)	266 545	39
TIMSS-Repeat	47,0* (13,5)	35,9* (12,9)	180 544	38
Beide	42,7* (9,8)	35,9* (8,3)	447 089	77
Beide, mit regionalen Dummies	28,6** (13,2)	41,7* (10,8)	447 089	77

Anmerkungen: Schätzer des Zentralprüfungskoeffizienten. Abhängige Variable: TIMSS-Testergebnis. Kontrollvariablen: 48 Schüler-, Familien-, Ressourcen-, Lehrer- und Institutionenmerkmale. Cluster-robuste Standardfehler (auf Länderebene) in Klammern.

Signifikanzniveau (basierend auf Cluster-robusten Standardfehlern): \* 1 %, \*\* 5 %.

Quelle: Wößmann (2002b).

fasst.<sup>5</sup> Der Schätzfehler hat mehrere Komponenten auf verschiedenen Ebenen:  $\mu$  ist eine länderspezifische,  $\eta$  eine schulspezifische,  $\nu$  eine klassenspezifische und  $\varepsilon$  eine schüler-spezifische Komponente.<sup>6</sup>

Die in Tabelle 1 berichteten Schätzergebnisse für den Zentralprüfungseffekt belegen, dass Schüler in Schulsystemen mit zentralen Prüfungen statistisch signifikant besser abschneiden als Schüler in Schulsystemen ohne zentrale Prüfungen. Dies gilt sowohl für Mathematik als auch für Naturwissenschaften, und es gilt sowohl für TIMSS als auch für TIMSS-Repeat. Im gepoolten Datensatz beträgt der Zentralprüfungseffekt 42,7 % einer internationalen Standardabweichung in Mathematik und 35,9 % in Naturwissenschaften. Dies entspricht in etwa dem Leistungsunterschied zwischen Schülern der siebten Klasse und Schülern der achten Klasse, also dem in einem ganzen Schuljahr erlernten Wissen. Die Ergebnisse von TIMSS-Repeat untermauern demzufolge zuvor ausschließlich anhand von TIMSS-95 erzielte Ergebnisse (Bishop 1997, Wößmann 2002a), dass Schülerleistungen in Zentralprüfungssystemen höher sind. Auch ist die Größe des für TIMSS-Repeat geschätzten Effektes statistisch nicht signifikant von dem für TIMSS-95 geschätzten Effekt zu unterscheiden.

Im Prinzip ist es denkbar, dass diese Kleinstquadrateschätzer des Zentralprüfungseffektes durch Endogenitätsprobleme verzerrt sind (vgl. etwa Büchel et al. 2003). So könnte es unberücksichtigte Faktoren („omitted variables“) auf Länderebene geben, die mit der Existenz von Zentralprüfungssystemen korrelieren und die die Korrelation mit den Schülerleistungen verursachen. Denkbar sind hier vor allem vier Bereiche: erstens andere institutionelle Gegebenheiten des Schulsystems, zweitens der allgemeine Zentralisationsgrad eines Landes, drittens die Homogenität der Bevölkerung und viertens kulturelle Unterschiede zwischen Ländern. Insoweit die Existenz von Zentralprüfungen nicht zufällig über die Länder verteilt

<sup>5</sup> Die einzelnen Kontrollvariablen finden sich in Tabelle A1 in Wößmann (2002b).

<sup>6</sup> Die Fehlerkomponenten werden durch so genannte Cluster-robuste Schätzungen (clustering-robust linear regression, CRLR) implementiert, die bei der Berechnung des Zentralprüfungseffektes mögliche Interdependenzen der Fehlerterme von Schülern innerhalb eines Landes – und, weiter unten, bei der Berechnung der Autonomieeffekte innerhalb einzelner Schulen – bei der Berechnung der Standardfehler berücksichtigen (vgl. Moulton 1986; Deaton 1997: 74–78). Der stratifizierten Stichprobenerhebung in TIMSS wird dadurch Rechnung getragen, dass die Observation eines jeden Schülers innerhalb seines Landes mit seiner Erhebungswahrscheinlichkeit gewichtet wird (vgl. DuMouchel und Duncan 1983, Wooldridge 2001); gleichzeitig werden alle Länder gleich gewichtet.

ist, sondern in zentralisierten oder homogenen Ländern besonders häufig auftritt oder mit anderen Institutionen oder kulturellen Merkmalen einhergeht und insoweit diese anderen Ländermerkmale Unterschiede in den Schülerleistungen hervorrufen, würde der einfache Kleinstquadrateschätzer des Zentralprüfungseffektes von solchen Einflüssen verzerrt sein.

In den ersten drei Bereichen ist es möglich, die Verzerrung zumindest zum größten Teil durch die Berücksichtigung zusätzlicher entsprechender Kontrollvariablen zu beheben. So beinhaltet die berichtete Spezifikation bereits eine große Anzahl institutioneller Kontrollvariablen (einschließlich der Zentralisierung des Curriculums und der Lehrbuchgenehmigung), und entsprechende Tests zeigen, dass ihre Berücksichtigung keinen signifikanten Einfluss auf den geschätzten Zentralprüfungseffekt hat. Zur Kontrolle für die allgemeine Zentralisierung des Systems und für die Bevölkerungshomogenität wurden zusätzlich der von der Zentralregierung kontrollierte Anteil an den Bildungsausgaben sowie ein Maß der ethnolinguistischen Fraktionalisierung der Bevölkerung als Kontrollvariablen in die Spezifikation eingeführt, ohne dass sich dadurch der geschätzte Zentralprüfungseffekt nennenswert geändert hätte.

Um schließlich zu testen, ob der Zentralprüfungseffekt sonstige kulturelle Unterschiede zwischen Ländern auffängt, können regionale (kontinentale) Dummies als zusätzliche Kontrollvariablen eingefügt werden. Dadurch erfolgt die Schätzung des Zentralprüfungseffektes ausschließlich aufgrund von innerregionaler Variation. So beeinflussen interregionale kulturelle Unterschiede wie etwa zwischen asiatischen und europäischen Wertesystemen den Zentralprüfungsschätzer nicht mehr. Wie in Tabelle 1 berichtet, ergeben die Schätzungen sogar dann statistisch signifikante Zentralprüfungseffekte, wenn sämtliche Variation zwischen den neun Regionen Westeuropa, Osteuropa, Nordamerika, Südamerika, Ozeanien, Asien, Mittlerer Osten, Nordafrika und Südafrika unberücksichtigt bleibt. Der geschätzte Zentralprüfungseffekt scheint also weder durch andere institutionelle Unterschiede noch durch die allgemeine Zentralisation oder Homogenität eines Landes noch durch kulturelle Unterschiede bedingt zu sein, sondern einen Effekt der externen Prüfungen auf die Schülerleistungen wiederzugeben.

### 3.3 Schulautonomie und Schülerleistungen mit und ohne zentrale Prüfungen

Um nun zu untersuchen, ob sich – wie zuvor hergeleitet – die Existenz von Zentralprüfungen auf den Einfluss von Schulautonomie auf erzielte Schülerleistungen auswirkt, füge ich in die Schätzgleichung (1) zusätzlich Interaktionsterme zwischen Zentralprüfungen  $Z$  und den Indikatoren der Schulautonomie  $A$  ein:

$$T_{iksl} = \alpha Z_l + A_{ksl} \beta_1 + (Z_l A_{ksl}) \beta_2 + C_{iksl} \gamma + \mu_l + \eta_s + \nu_k + \varepsilon_i. \quad (2)$$

Die geschätzten Interaktionseffekte geben an, ob sich der Einfluss von Schulautonomie in verschiedenen Entscheidungsbereichen zwischen Systemen mit und ohne Zentralprüfungen unterscheidet. Die kompletten Schätzergebnisse für die Autonomie- und Interaktionseffekte werden in Tabelle A1 im Anhang berichtet.<sup>7</sup> Die markantesten Befunde sollen im Folgenden einzeln anhand von Schaubildern diskutiert werden.

<sup>7</sup> Die hier berichteten Ergebnisse beziehen sich auf die Mathematikleistungen; ähnliche Ergebnisse ergeben sich für Naturwissenschaftsleistungen (Wößmann 2002b).

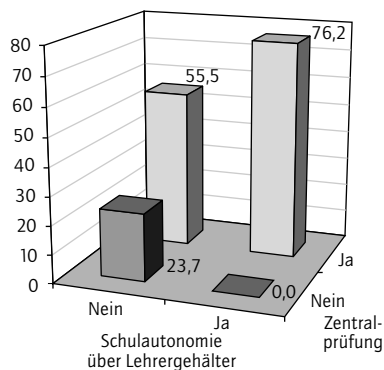
Die folgenden Abbildungen stellen, jeweils für verschiedene Entscheidungsbereiche, die Schülerleistungen unter den vier Zuständen dar, die sich aus dem Vorhandensein und Nichtvorhandensein von Schulautonomie und Zentralprüfungen ergeben: die Leistungen von Schülern in Schulen ohne Autonomie in Systemen ohne Zentralprüfungen, mit Autonomie ohne Zentralprüfungen, ohne Autonomie mit Zentralprüfungen und sowohl mit Autonomie als auch mit Zentralprüfungen. Die Schülerleistungen werden in jeder Abbildung relativ zum Zustand mit dem niedrigsten Leistungsstand angegeben.<sup>8</sup>

Abbildung 2a stellt den Fall dar, ob Schulen Verantwortung bei der Entscheidung über Lehrergehälter haben. In Systemen ohne Zentralprüfungen hat Schulautonomie über Lehrergehälter einen *negativen* Effekt auf die Schülerleistungen. In Systemen mit Zentralprüfungen ist die Schülerleistung – entsprechend den zuvor berichteten Ergebnissen – zunächst einmal höher als in Systemen ohne Zentralprüfungen, sowohl im Fall mit als auch im Fall ohne Schulautonomie. Zusätzlich sticht aber noch hervor, dass sich der Einfluss von Schulautonomie auf die Schülerleistungen in Systemen mit Zentralprüfungen komplett umdreht: Die Gehaltsautonomie der Schulen hat in Zentralprüfungssystemen *positive* Auswirkungen auf die Schülerleistungen.<sup>9</sup>

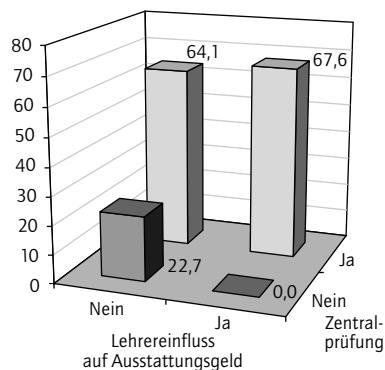
Abbildung 2

**Zentralprüfungen und Autonomie bei Opportunismus und lokalem Wissensvorsprung**  
TIMSS-Mathematik-Testergebnis (relativ zur niedrigsten Kategorie)

(a) Schulautonomie über Lehrergehälter



(b) Lehreinfluss auf Ausstattungsgeld



Quelle: Wößmann (2002b).

Damit stellen sich Entscheidungen über Lehrergehälter als ein Entscheidungsbereich dar, in dem sowohl Anreize für opportunistisches Verhalten als auch lokale Wissensvorteile vorzuliegen scheinen (vgl. Abbildung 1): Ohne zentrale Prüfungen dominieren die negati-

<sup>8</sup> Die Schätzungen, auf denen die Abbildungen basieren, kontrollieren jeweils für die gesamten Familien-, Ressourcen- und Institutionenkontrollvariablen der Schätzgleichung (2), aber – im Gegensatz zu den in Tabelle A1 berichteten Ergebnissen – nicht für weitere Interaktionseffekte zwischen Zentralprüfungen und Familien- und Institutionsvariablen, da ansonsten der eigene Effekt der Zentralprüfungen nur recht unpräzise geschätzt würde und die Balken für Effekte in Zentralprüfungssystemen folglich auf unpräzisen Schätzern basieren müssten (Wößmann 2002b).

<sup>9</sup> Soweit nicht anders berichtet, sind alle in den Abbildungen 2 bis 4 abgebildeten Effekte statistisch signifikant.

ven Leistungswirkungen opportunistischer Entscheidungen der Schulen, da dieses dezentrale opportunistische Verhalten extern nicht festgestellt und deshalb nicht sanktioniert werden kann. Schulische Entscheidungsträger sehen sich also nicht genötigt, die Lehrergehälter auf eine Weise festzusetzen, die zur Leistungssteigerung der Schüler beitragen könnte, sondern können diese Entscheidungsautonomie zur Förderung anderweitiger Interessen einsetzen. Demgegenüber schaffen Zentralprüfungen Informationen darüber, ob die Schulen gute Leistungsergebnisse erzielen oder nicht, so dass Aufsichtsbehörden und Eltern mögliche Konsequenzen aus leistungsminderndem Verhalten der Schulen ziehen können. Dies schafft Anreize für die Entscheidungsträger in den Schulen, ihre Autonomie über Lehrergehälter nicht opportunistisch auszunutzen, sondern effektiv zur Leistungsförderung der Schüler einzusetzen. Dabei kommen nun Vorteile lokalen Wissens zum Tragen, da schulische Entscheidungsträger besser als eine zentrale Instanz wissen dürften, welcher Lehrer eine Belohnung für gute Arbeit verdient.

Ähnlich liegt der Fall, wenn es darum geht, ob Ausstattungsentscheidungen dergestalt dezentralisiert sind, dass Lehrer Einfluss auf das für die Ausstattung verfügbare Geld haben (Abbildung 2b). Auch in diesem Entscheidungsbereich wirkt sich eine dezentrale Entscheidungsautonomie in Systemen ohne Zentralprüfungen negativ auf die erzielten Schülerleistungen aus, während sie sich in Systemen mit Zentralprüfungen positiv auswirkt. Allerdings ist dieser Unterschied zwischen Schulen mit und ohne Lehrereinfluss auf die finanzielle Ausstattung in Systemen mit Zentralprüfungen nicht statistisch signifikant. Dies könnte entweder daran liegen, dass in diesem Fall opportunistisches Verhalten durch die Zentralprüfungen nicht gänzlich unterbunden wird und dieses deshalb die positiven Effekte lokalen Wissens abschwächt, oder daran, dass in diesem Fall kein signifikanter lokaler Wissensvorsprung vorliegt.

Das Gleiche scheint in verstärktem Maße in den in Abbildung 3 dargestellten Entscheidungsbereichen der Fall zu sein. In Systemen ohne Zentralprüfungen hat Schulautonomie über die Schulbudgets einen negativen Einfluss auf die Schülerleistungen (Abbildung 3a), was auf Anreize für opportunistisches Verhalten in finanziellen Bereichen zurückzuführen sein dürfte. In Systemen mit Zentralprüfungen verschwindet dieser negative Einfluss der Schulautonomie, ohne sich aber in einen signifikanten positiven Einfluss umzukehren. Dies könnte daran liegen, dass externe Instanzen durchaus keinen Wissensnachteil in Budgetfragen haben müssen im Vergleich zu einzelnen Schulen, denen oftmals das nötige Fachpersonal fehlt.

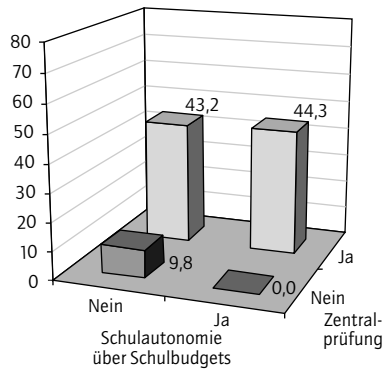
Das gleiche Bild ergibt sich, wenn es darum geht, ob die Fachlehrer der Schulen kollektiv einen Einfluss auf das zu lehrende Curriculum haben (Abbildung 3b). Ohne das Monitoring zentraler Prüfungen hat ein solcher kollektiver Lehrereinfluss einen negativen Effekt auf die erzielten Schülerleistungen, was auf opportunistische Interessen des Lehrerkollektivs in Bezug auf die zu leistende Arbeitslast zurückgeführt werden könnte. Ist ein zentrales Prüfungssystem gegeben, so schwächt sich dieser negative Leistungseffekt in einen insignifikanten Effekt der Lehrerautonomie ab. Dies könnte man im obigen Modellrahmen mit einem abgeschwächten opportunistischen Verhalten bei gleichzeitig für kollektive Lehrerentscheidungen nicht vorliegende lokale Wissensvorsprünge rationalisieren.<sup>10</sup>

**10** Man beachte, dass in der zugrunde liegenden Schätzung für den individuellen Lehrereinfluss auf das Curriculum kontrolliert wird, so dass individuelle Wissensvorteile der Lehrer konstant gehalten sind.

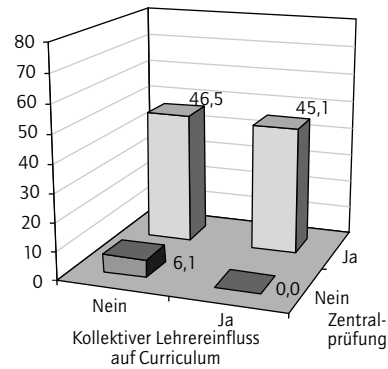
Abbildung 3

**Zentralprüfungen und Autonomie bei Opportunismus ohne lokalen Wissensvorsprung**  
TIMSS-Mathematik-Testergebnis (relativ zur niedrigsten Kategorie)

(a) Schulautonomie über Schulbudgets



(b) Kollektiver Lehrereinfluss auf Curriculum



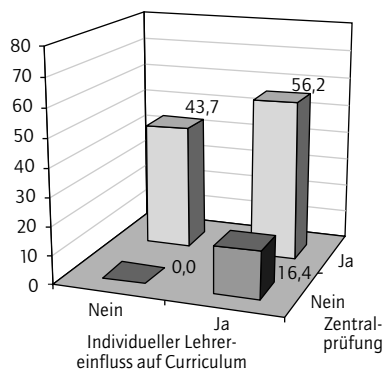
Quelle: Wößmann (2002b).

Anders sieht es aus, wenn nicht die Lehrer kollektiv, sondern die einzelnen Lehrer individuell Einflussmöglichkeiten auf das Curriculum haben (Abbildung 4a). In diesem Fall ergibt sich in Systemen sowohl mit als auch ohne zentrale Prüfungen ein positiver Effekt der Lehrerautonomie auf die erzielten Schülerleistungen, der sich statistisch nicht signifikant zwischen den beiden Systemen unterscheidet. Im betrachteten Modellrahmen hieße das, dass die individuellen Lehrer in diesem Fall nicht die Möglichkeit haben, zusätzlich zur kollektiven Einflussnahme noch opportunistische Verhaltensweisen durchzusetzen, während sie aber individuell lokale Wissensvorteile besitzen.

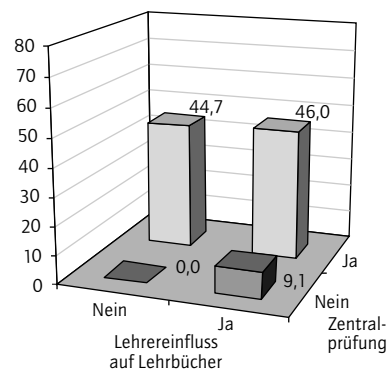
Abbildung 4

**Zentralprüfungen und Autonomie bei lokalem Wissensvorsprung ohne Opportunismus**  
TIMSS-Mathematik-Testergebnis (relativ zur niedrigsten Kategorie)

(a) Individueller Lehrereinfluss auf Curriculum



(b) Lehrereinfluss auf Lehrbücher



Quelle: Wößmann (2002b).

In Abbildung 4b ist schließlich ein Entscheidungsbereich abgebildet, in dem sich das Vorhandensein zentraler Prüfungen abschwächend auf einen positiven Autonomieeffekt auswirkt: Hat der einzelne Lehrer in Systemen ohne Zentralprüfungen einen Einfluss auf das verwendete Lehrbuch, so wirkt sich das positiv auf die erzielten Schülerleistungen aus. Dies dürfte an lokalen Wissensvorsprüngen liegen, die nicht durch opportunistische Interessenlagen konterkariert werden, da sich der Lehrer durch die Wahl unpassender Lehrbücher wohl eher selbst schaden würde. Dieser positive Effekt von Lehrerautonomie wird in Systemen mit Zentralprüfungen kleiner und statistisch insignifikant. Im Gegensatz zur vereinfachenden Darstellung von Abbildung 1 kann sich also bei lokalem Wissensvorsprung trotz fehlender Anreize für opportunistisches Verhalten ein Unterschied in der Auswirkung von Autonomie auf Schülerleistungen zwischen Systemen mit und ohne Zentralprüfungen ergeben. Diese Abschwächung des Autonomieeffektes könnte darauf zurückzuführen sein, dass das Vorhandensein zentraler Prüfungen die dezentralen Entscheidungsträger daran hindert, ihre lokalen Wissensvorteile voll zum Tragen kommen zu lassen. Es sollte aber beachtet werden, dass auch in diesem Fall die Gesamtleistungen von Schülern in Zentralprüfungssystemen immer noch weit besser sind als in Systemen ohne Zentralprüfungen.

Der letzte in Abbildung 1 abgebildete Fall, in dem es weder Anreize für opportunistisches Verhalten noch lokale Wissensvorsprünge gibt, ist für die Betrachtung von Autonomieeffekten relativ uninteressant, da Autonomie hier weder mit noch ohne Zentralprüfungen einen Einfluss auf die erzielten Schülerleistungen hat. Dies könnte für Mathematik in dem in Tabelle A1 berichteten Entscheidungsbereich der Lehrerautonomie bei Entscheidungen über die Art der Ausstattung der Fall sein, wo sich keine signifikanten Autonomieeffekte ergeben. Allerdings ergibt sich für Naturwissenschaften ein kleiner, aber statistisch signifikanter positiver Autonomieeffekt, der in Zentralprüfungssystemen leicht abgeschwächt wird, was auf das Vorhandensein lokaler Wissensvorsprünge ohne Opportunismus hindeuten würde. Tabelle A1 berichtet das gleiche Schema für Schulautonomie beim Einkauf von Ausstattungsmitteln. Lehrerautonomie bei der Auswahl der Themengebiete ist demgegenüber ein weiteres Beispiel für einen Entscheidungsbereich mit Anreizen für Opportunismus und ohne lokalen Wissensvorsprung, bei dem ein negativer Autonomieeffekt in Systemen ohne Zentralprüfungen durch Zentralprüfungen weitgehend ausgeschaltet wird.<sup>11</sup>

#### 4 Schlussbemerkungen

Indem zentrale Prüfungen helfen, Informationsasymmetrien im Bildungssystem zu überwinden, tragen sie dazu bei, dass ansonsten dezentral organisierte Schulsysteme die Vorteile dezentralen Wissens zum Tragen bringen können, ohne dass diese durch lokales opportunistisches Verhalten konterkariert würden. So können Zentralprüfungen quasi wie eine „Währung“ des Bildungssystems wirken: Als zentral bereitgestelltes Leistungs- und Bewertungsmaß senken sie Transaktionskosten, die aufgrund unvollständiger Verträge bei asymmetrischer Information zwischen Prinzipalen und Agenten entstehen. Damit erweisen sich zentrale Prüfungen als komplementär zur Schulautonomie: Sie verringern oppor-

<sup>11</sup> Die schwer zu rationalisierenden Ergebnisse für den Fall der Schulautonomie bei der Lehrerauswahl in Tabelle A1 sind auf den Mathematik-Fall beschränkt; in Naturwissenschaften ergibt sich in Systemen ohne Zentralprüfungen kein Autonomieeffekt, der sich mit Zentralprüfungen in einen positiven Effekt umwandelt. Dies ist der einzige Fall, in dem die Naturwissenschaftsergebnisse grundlegend von denen in Mathematik abweichen. Zur Diskussion weiterer institutioneller Interaktionseffekte mit Zentralprüfungen vgl. Wößmann (2002b).

tunistisches Verhalten dezentraler Entscheidungsträger und werden so zur Voraussetzung dafür, dass dezentrale Schulsysteme funktionsfähig sind.

Zentrale Prüfungssysteme bewirken also nicht nur Verhaltensänderungen bei den Schülern, sondern auch bei allen anderen am Bildungsprozess beteiligten Akteuren. Sie verändern die Anreize und richten sie auf eine Förderung der schulischen Leistungen aus. Die Änderungen im Verhalten der Entscheidungsträger in den Schulen und damit in der Wirkung von Schulautonomie erweisen sich also als ein Mechanismus, durch den Zentralprüfungen zu besseren Schülerleistungen führen. Deshalb würde eine effiziente Bildungspolitik so aussehen, dass sie zentrale Prüfungen mit Schulautonomie verbindet, das heißt, sie würde Standards vorgeben und ihr Erreichen überprüfen, aber es gleichzeitig den Schulen überlassen, wie diese Standards zu erreichen sind.

Für die Währung ist des Öfteren festgestellt worden, dass vor allem die ärmeren Teile einer Gesellschaft getroffen werden, wenn die Funktion des Geldes als Wertmaßstab – etwa durch Inflation – schwindet. In ähnlicher Weise lässt sich fragen, ob auch Zentralprüfungen Verteilungswirkungen aufweisen. Detailliertere Schätzungen, die Interaktionseffekte zwischen Zentralprüfungen und den Indikatoren des familiären Hintergrunds der Schüler berücksichtigen, zeigen, dass der Effekt von Zentralprüfungen tatsächlich unterschiedlich groß ist für Schüler mit unterschiedlichem familiären Hintergrund (Wößmann 2002b). So ist der Effekt des Ausbildungsstandes der Eltern auf die Leistungen ihrer Kinder in Systemen ohne Zentralprüfungen weit größer als in Systemen mit Zentralprüfungen. Dieser Unterschied zwischen den beiden Systemen ist, vor allem in Mathematik, statistisch hoch signifikant. Ebenso ist der Leistungsrückstand von Immigrantenkindern der ersten und zweiten Generation in Systemen mit Zentralprüfungen weit geringer als in Systemen ohne Zentralprüfungen. Diese Befunde legen nahe, dass zentrale Prüfungen zu ausgeglicheneren Bildungschancen für Schüler mit unterschiedlichem familiärem Hintergrund beitragen – oder anders ausgedrückt: Auch im Bildungssystem hat das Fehlen der „Währung“ Zentralprüfung negative Verteilungswirkungen.

Zentrale Prüfungen müssen nicht jegliche Art der Profilbildung von Schulen verhindern. Bei den hier untersuchten Leistungswirkungen geht es um Basiskompetenzen, die 13-jährige Schüler in Mathematik aufweisen sollten. Befürworter eher homogener Schulsysteme und Befürworter eher differenzierter Schulsysteme sollten sich darauf einigen können, dass eine Beherrschung solcher Basiskompetenzen ein zentrales Ziel eines jeden Schulsystems darstellen sollte. In dieser Hinsicht ist es auch von Bedeutung, dass es sich bei dem in den berichteten Analysen verwendeten Leistungsmaß, den TIMSS-Testergebnissen, nicht um die jeweilige zentrale Prüfung eines Landes handelt, sondern um einen unabhängigen Leistungstest, der von allen beteiligten Ländern als dem jeweiligen Mathematik- und Naturwissenschaftscurriculum entsprechend akzeptiert wurde. Das bedeutet, dass Zentralprüfungen nicht einfach nur dazu führen, dass die Lehrer ihre Schüler lediglich mögliche Fragen der jeweiligen Zentralprüfung auswendig lernen lassen und dass die Schüler den entsprechenden Test „einpacken“ – denn dies sollte ihre Leistungen in den TIMSS-Tests nicht beeinflussen. Stattdessen legen die berichteten Schätzergebnisse nahe, dass Zentralprüfungen tatsächlich zu einem besseren mathematischen und naturwissenschaftlichen Basiswissen der Schüler führen. Darüber hinaus sei noch darauf verwiesen, dass es auch bei externen Prüfungen durchaus möglich ist, die Leistungsanforderungen zu differenzieren und Schwerpunktsetzungen der Schulen abzubilden.

Über die berichteten empirischen Ergebnisse, die sich auf zentrale Abschlussprüfungen am Ende der Sekundarbildung beziehen, hinaus legt es die vorgestellte Theorie nahe, dass regelmäßige Zentralprüfungen im Verlaufe der Primar- und Sekundarstufe durchaus noch zu weiteren positiven Effekten führen könnten. Solche regelmäßigen externen Prüfungen in verschiedenen Jahrgangsstufen würden die Informationslage im Bildungssystem noch weiter verbessern, und durch ihre frühzeitige Verfügbarkeit würden sie es auch ermöglichen, bei unbefriedigenden Leistungen schon weit vor Abschluss der Sekundarbildung gegenzulenken.

## Literaturverzeichnis

- Bishop, John H. (1997): The Effect of National Standards and Curriculum-Based Exams on Achievement. *American Economic Review*, 87 (2), 260–264.
- Bishop, John H. und Ludger Wößmann (2003): Institutional Effects in a Simple Model of Educational Production. *Education Economics* (im Erscheinen).
- Brunner, Karl und Allan H. Meltzer (1971): The Use of Money: Money in the Theory of an Exchange Economy. *American Economic Review*, 61 (5), 784–805.
- Büchel, Felix, Hendrik Jürges und Kerstin Schneider (2003): Die Auswirkungen zentraler Abschlussprüfungen auf die Schulleistung – Quasi-experimentelle Befunde aus der deutschen TIMSS-Stichprobe. In diesem Heft, 238–251.
- Deaton, Angus (1997): *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore, The Johns Hopkins University Press.
- DuMouchel, William H. und Greg J. Duncan (1983): Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples. *Journal of the American Statistical Association*, 78 (383), 535–543.
- Gonzalez, Eugenio J. und Julie A. Miles (Hrsg.) (2001): *User Guide for the TIMSS 1999 International Database*. Chestnut Hill, MA, Boston College.
- Gonzalez, Eugenio J. und Teresa A. Smith (Hrsg.) (1997): *User Guide for the TIMSS International Database: Primary and Middle School Years*. Chestnut Hill, MA, Boston College.
- Gundlach, Erich, Ludger Wößmann und Jens Gmelin (2001): The Decline of Schooling Productivity in OECD Countries. *Economic Journal*, 111 (471), C135–C147.
- Gundlach, Erich, Jose Navarro de Pablo und Natascha Weisert (2003): Education Is Good for the Poor: A Note on Dollar and Kraay. In: Anthony Shorrocks und Rolph van der Hoeven (Hrsg.): *Growth, Inequality and Poverty*. Oxford, Oxford University Press (im Erscheinen).
- Hanushek, Eric A. (2002): Publicly Provided Education. In: Auerbach, Alan und Martin Feldstein (Hrsg.): *Handbook of Public Economics*, Vol. 4. Amsterdam, Elsevier, 2045–2141.
- Hanushek, Eric A. (2003): The Failure of Input-Based Schooling Policies. *Economic Journal*, 113 (485), F64–F98.
- Hayek, Friedrich August (1945): The Use of Knowledge in Society. *American Economic Review*, 35 (4), 519–530.
- Issing, Ottmar (1998): *Einführung in die Geldtheorie*. 11. Aufl. München, Vahlen.
- Laffont, Jean-Jacques und David Martimort (2002): *The Theory of Incentives: The Principal-Agent Model*. Princeton, Princeton University Press.
- Moulton, Brent R. (1986): Random Group Effects and the Precision of Regression Estimates. *Journal of Econometrics*, 32 (3), 385–397.
- Schumpeter, Joseph A. (1970): *Das Wesen des Geldes*. Göttingen, Vandenhoeck & Ruprecht.



- Summers, Anita A. und Amy W. Johnson (1996): The Effects of School-Based Management Plans. In: Hanushek, Eric. A. und Dale W. Jorgenson (Hrsg.): *Improving America's Schools: The Role of Incentives*. Washington, D. C., National Academy Press, 75–96.
- Weiß, Manfred (1998): Schulautonomie im Licht mikroökonomischer Bildungsforschung. In: Robert K. von Weizsäcker (Hrsg.): *Deregulierung und Finanzierung des Bildungswesens*. Schriften des Vereins für Socialpolitik, Bd. 262. Berlin, Duncker & Humblot, 15–47.
- Wooldridge, Jeffrey M. (2001): Asymptotic Properties of Weighted *M*-Estimators for Standard Stratified Samples. *Econometric Theory*, 17 (2), 451–470.
- World Bank (1999): *World Development Report 1998/99: Knowledge for Development*. Oxford, Oxford University Press.
- Wößmann, Ludger (2002a): *Schooling and the Quality of Human Capital*. Berlin, Springer.
- Wößmann, Ludger (2002b): *Central Exams Improve Educational Performance: International Evidence*. Institut für Weltwirtschaft, Kieler Diskussionsbeiträge 397. Kiel.
- Wößmann, Ludger (2003a): Specifying Human Capital: A Review and Some Extensions. *Journal of Economic Surveys*, 17 (3), 239–270.
- Wößmann, Ludger (2003b): Schooling Resources, Educational Institutions, and Student Performance: The International Evidence. *Oxford Bulletin of Economics and Statistics*, 65 (2), 117–170.
- Wößmann, Ludger und Martin R. West (2002): *Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS*. Harvard University, Program on Education Policy and Governance Research Paper, PEPG/02-02.

## Anhang

Tabelle A1

### Interaktionseffekte von Zentralprüfungen und Schulautonomie

	In Systemen ohne Zentralprüfungen <sup>1</sup>	Änderung in Systemen mit Zentralprüfungen <sup>2</sup>
<b>Schulautonomie</b>		
Schulbudget	-6,9** (2,8)	7,7** (3,5)
Ausstattungsinkauf	7,1** (3,2)	-5,7 (5,0)
Lehrerauswahl	21,6* (2,6)	-20,2* (3,1)
Lehrergehälter	-28,3* (3,6)	50,2* (4,1)
<b>Lehrereinfluss</b>		
Gelder für Ausstattung	-24,7* (5,1)	29,1* (6,3)
Art der Ausstattung	3,0 (2,8)	-3,5 (3,8)
Themengebiete	-12,3* (2,3)	8,7* (2,8)
Lehrbuch	11,6* (3,1)	-11,7* (3,6)
Curriculum		
Individuelle Lehrer	14,6* (2,1)	-3,9 (2,7)
Fachlehrer kollektiv	-5,0** (2,4)	2,8 (3,1)
Alle Lehrer kollektiv	-14,7* (2,1)	6,5* (2,8)
Lehrergewerkschaften	-8,5 (5,4)	-29,5* (8,7)
Schüler (Observationseinheit)	447 089	
Schulen (Primäre Stichprobeneinheit)	12 175	
Länder	77	
R <sup>2</sup>	0,296	

Anmerkungen: Schätzer des jeweiligen Autonomiekoeffizienten und des Interaktionskoeffizienten mit Zentralprüfungen. Abhängige Variable: TIMSS-Mathematiktestergebnis. Kontrollvariablen: 36 Schüler-, Familien-, Ressourcen-, Lehrer- und Institutionenmerkmale sowie 17 Interaktionseffekte von Schüler-, Familien- und Institutionenmerkmalen mit Zentralprüfungen. Cluster-robuste Standardfehler (auf Schulebene) in Klammern.

Signifikanzniveau (basierend auf Cluster-robusten Standardfehlern): \* 1 %, \*\* 5 %.

**1** Koeffizient des Schätzers an der jeweiligen Autonomievariablen ( $\beta_1$  in Schätzgleichung [2]).

**2** Koeffizient des Schätzers am Interaktionsterm zwischen Zentralprüfungen und der jeweiligen Autonomievariablen ( $\beta_2$  in Schätzgleichung [2]).

Quelle: Wößmann (2002b).