

(vgl. Art. 91b Abs. 1 S. 1 Nr. 1 GG), als hinderlich erweisen, muss sich zeigen. Da der Bund fortan nur dann Finanzhilfen an die Länder gewähren darf, wenn er auch für die zu fördernde Materie zuständig ist, wurde das Einfallstor für eine sachliche Beeinflussung der Länder durch den Bund an der grundgesetzlichen Kompetenzordnung vorbei ein Stück weit geschlossen. Nicht übersehen werden darf aber, dass dem Bund weit reichende Gesetzgebungskompetenzen namentlich im Hochschulbereich zustehen, was die Möglichkeit entsprechender Finanzhilfen nach sich zieht.

*Verf.: Dr. Ferdinand Wollenschläger, Akademischer Rat, Lehrstuhl für Öffentliches Recht und Staatsphilosophie, Juristische Fakultät, Ludwig-Maximilians-Universität München, 80539 München*

*Manfred Weiß/Johannes Bellmann*

## **Bildungsfinanzierung in Deutschland und Schulqualität – eine gefährdete Balance?\***

### **1** Problemhintergrund

Das Zusammentreffen von anhaltender staatlicher Finanzkrise und weiter steigenden Anforderungen an das Bildungssystem hat dort zu einer beträchtlichen Verschärfung der finanziellen Engpasssituation geführt.

Verstärkte Aufmerksamkeit findet deshalb in Politik und Wissenschaft die Frage, wie dem Bildungsbereich gleichwohl eine angemessene Ressourcenausstattung gesichert werden kann. Dabei müssen neue Begründungsmuster und Strategien gefunden werden. Verbraucht hat sich offensichtlich das rituelle Lamento, der Bildungssektor zähle zu den besonders Not leidenden öffentlichen Aufgabenbereichen und müsse deshalb im gesamtstaatlichen Mittelverteilungsprozess eine Besserstellung erfahren. Auch die in der politischen Rhetorik viel bemühte These, Bildung sei die Hauptproduktivkraft einer wissensbasierten Ökonomie, ist bisher ohne sichtbaren Effekt für die Ressourcenmobilisierung geblieben.

Die für Deutschland außerordentlich ernüchternden Ergebnisse internationaler Schulleistungsstudien im Sekundarbereich (TIMSS, PISA) haben zwar eine breit angelegte „Qualitäts offensive“ ausgelöst, die eine ganze Reihe budgetwirksamer Maßnahmen vorsieht; doch stehen diese allesamt unter dem Vorbehalt der Finanzierbarkeit. Immerhin liegen mittlerweile zwei auf ausformulierten Reformvorschlägen basierende Schätzungen des Ressourcenbedarfs vor, an denen sich die Mittelforderungen der Bildungsseite orientieren können. Nach Berechnungen der Vereinigung der Bayerischen Wirtschaft (2004) würde das von ihr unter Mitwirkung zahlreicher Experten erarbeitete Reformkonzept (vgl. im Einzelnen Vereinigung der Bayerischen Wirtschaft 2003) einen jährlichen Mehrbedarf an Mitteln von 35 Mrd. Euro erfordern, was einer Steigerung der aktuellen Bildungsausgaben um etwa 30 % entspricht. Drei Viertel der Mehr-

---

\* Der vorliegende Beitrag erscheint zeitgleich in *van Buer, J./Wagner, C.* (Hrsg.), *Qualität von Schule – Entwicklungen zwischen erweiterter Selbstständigkeit, definierten Bildungsstandards und strikter Ergebniskontrolle*, Frankfurt a.M. u.a. 2007.

aufwendungen entfielen auf die Länder und Gemeinden, die die Hauptlast der öffentlichen Bildungsfinanzierung tragen. Bescheidener fällt der von Klemm (2005) ermittelte reforminduzierte Mehrbedarf an Finanzmitteln aus. Unter Berücksichtigung demographisch bedingter Minder Ausgaben müsste danach 2015 das öffentliche Bildungsbudget lediglich um 10 Mrd. Euro (in Preisen von 2002) aufgestockt werden.

Zur Frage, wie der zusätzliche Mittelbedarf finanziert werden soll, finden sich in den Modellrechnungen indes – wie seinerzeit auch im Bildungsgesamtplan und seiner (gescheiterten) Fortschreibung – keine konkreten Hinweise. Beide Projektionen stimmen darin überein, dass die öffentlichen Haushalte die primäre Finanzierungsquelle darstellen. Dass sich diese in einem prekären Zustand befinden, ist weithin bekannt. Selbst bei einer nachhaltigen wirtschaftlichen Erholung dürfte es nicht gelingen, den für Qualitätsverbesserungen im Bildungsbereich für notwendig erachteten zusätzlichen Mittelbedarf allein über die Zuweisung der anteiligen staatlichen Mehreinnahmen zu decken. Deshalb wird vielfach eine Neufestlegung der Prioritäten bei den öffentlichen Ausgaben zugunsten des Bildungssektors gefordert. Begründet wird dies nicht nur mit der besonderen strategischen Bedeutung von Bildung für Wirtschaft und Gesellschaft, sondern auch mit der rückläufigen Entwicklung des Anteils der Bildungsausgaben am Bruttoinlandsprodukt und der unterdurchschnittlichen Position Deutschlands bei diesem Indikator im internationalen Vergleich.<sup>1</sup>

Der mit durchaus überzeugenden Argumenten vorgetragene Forderung nach einer Erhöhung des BIP-Anteils für Bildung war indes schon in der Vergangenheit kein Erfolg beschieden: Sei Jahren liegen die Steigerungsraten der Bildungsausgaben unter denen der Wertschöpfung<sup>2</sup>. Ob es in Zukunft gelingen wird, den BIP-Anteil für Bildung auf dem Wege einer Neufestlegung der Prioritäten im gesamtstaatlichen Mittelverteilungsprozess anzuheben, ist angesichts des steigenden Mittelbedarfs anderer Bereiche höchst fraglich. Was die für die Bildungsfinanzierung besonders wichtigen Länderhaushalte betrifft, kann von einer weiteren Verschärfung der Engpasssituation ausgegangen werden. Die Berechnung von „Tragfähigkeitslücken“ verweist auf einen erheblichen Konsolidierungsbedarf. Bereits bis 2008 müssten danach in der Mehrzahl der Länder die Ausgaben (ohne Zinsausgaben) um 10 % gesenkt werden – in Berlin und Bremen sogar um mehr als 20 % –, um die Schuldenquote (die Verschuldung gemessen am BIP) konstant zu halten (vgl. Hauptmeier 2005). Zu berücksichtigen sind ferner die steigenden Kapitaldienst- und Versorgungszahlungen, die den politischen Handlungsspielraum der Länder längerfristig erheblich einengen werden. Nach Berechnungen des Bundesministeriums des Inneren (2005, 64) werden die Ausgaben der Länder für Pensionen bis 2030 um 70 % steigen. Welche Belastungen auf die Landeshaushalte bei unveränderter Finanzpolitik zukommen, veranschaulicht beispielhaft eine Modellrechnung für Niedersachsen (Homburg 2005): Danach wären 2030 54 % der Einnahmen allein durch Pensions- und Zinsausgaben gebunden; die aktuelle Quote liegt bei 24 %.

Die skizzierte Problemlage der öffentlichen Haushalte lässt die Annahme wenig realistisch erscheinen, der für intendierte Qualitätsverbesserungen im Bildungsbereich erforderliche Mehrbedarf an Ressourcen könne allein über Budgetzuwächse finanziert werden. In den Blick rücken deshalb zunehmend andere strategische Ansatzpunkte: Mittelumschichtungen innerhalb des

<sup>1</sup> Mit einem BIP-Anteil von 4,4 % lag Deutschland 2003 mit seinen öffentlichen Bildungsausgaben an 21. Stelle unter den 29 OECD-Staaten, die im Mittel 5,2 % ihrer Wertschöpfung für Bildung aufwendeten. Der deutlich unterdurchschnittliche öffentliche Anteil wird zwar z.T. korrigiert durch einen überdurchschnittlichen Anteil an privaten Bildungsausgaben (0,9 %), der vor allem von den Aufwendungen der Wirtschaft für den betrieblichen Teil der Ausbildung im Dualen System der Berufsbildung herrührt; dennoch bleibt auch der Gesamtanteil öffentlicher und privater Bildungsausgaben am BIP in Deutschland mit 5,3 % unter dem OECD-Mittel von 5,9 %.

<sup>2</sup> Konsortium Bildungsberichterstattung 2006, Kapitel B.

Bildungssektors, die verstärkte Nutzung privater Finanzierungsquellen und die Ausschöpfung von Effizienzreserven. Das Hauptaugenmerk gilt im Folgenden der letztgenannten Strategie. Die Ausführungen dazu beziehen sich auf die faktorielle Mittelverwendung im Schulbereich und eine Einschätzung des Effizienzverbesserungspotenzials neuer Finanzierungskonzepte.

## 2 Mittelumerschichtungen innerhalb des Bildungssektors

Auch die darauf bezogenen Empfehlungen finden ihre Begründung vielfach durch einen internationalen Vergleich einschlägiger Kennzahlen, denen der Status von Indikatoren der Systemqualität zugeschrieben wird. Die Orientierung der intrasektoralen Mittelallokation an internationalen Benchmarks (z.B. OECD-Mittel) dürfte indes kaum der Rationalitätsforderung genügen. Die meist dem Vergleich zugrunde liegenden teilnehmerbezogenen Bildungsaufwendungen z.B. reflektieren nicht nur unterschiedliche politische Prioritätensetzungen, sondern auch institutionelle Unterschiede sowie differente Nachfragebedingungen und Kostenstrukturen. Zudem ist die Validität solcher inputbezogenen Indikatoren der Outputqualität empirisch nicht zu sichern: Ein systematischer Zusammenhang z. B. mit Fachleistungen lässt sich nicht nachweisen (dazu im Einzelnen unten, Abschnitt 4.1). Eine evidenzbasierte Mittelverwendung, wie sie allenthalben gefordert wird, meint sicherlich etwas anderes: Die Mittel sind dort einzusetzen, wo sie den höchsten gesellschaftlichen Nettonutzen stiften oder den höchsten Grenzertrag je Geldeinheit bei der Leistungserstellung erbringen.

Im ersten Fall könnten *gesellschaftliche Ertragsraten*, wie sie die OECD für eine ganze Reihe von Mitgliedstaaten von Zeit zu Zeit berechnet, eine solche Orientierungsfunktion erfüllen. Dies propagiert auch der Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung in seinem Jahresgutachten 2004/05: „Die Verteilung öffentlicher Mittel auf die einzelnen Zweige des Bildungssystems muss sich an den sozialen Erträgen dieser Zweige orientieren“ (Sachverständigenrat 2004, 569). Allerdings werden durch die üblichen Renditeberechnungen die gesellschaftlichen Bildungserträge nur höchst unvollkommen erfasst, was deren Steuerungsfunktion für die Mittelallokation erheblich einschränkt.

Im zweiten Fall müssten Informationen darüber vorliegen, welchen Grenzbeitrag die Investitionen in den einzelnen Teilbereichen zum kumulativen Bildungsergebnis liefern. Darüber lässt sich bislang bestenfalls spekulieren. Auf die Bedeutung etwa des Kindergartenbesuchs für den Kompetenzerwerb (Lesen) im Primarbereich verweist die IGLU-Studie. Viertklässler, die einen Kindergarten besucht haben, sind ihren Mitschülern ohne Kindergartenerfahrung um ein halbes Schuljahr voraus (vgl. Bos et al. 2003). Belegt sind durch US-amerikanische Studien mit experimentellem Untersuchungsdesign schulische Erträge und zum Teil auch die damit verbundene Nettokostenersparnis von Interventionsmaßnahmen in der Vorschulerziehung für Risikogruppen (vgl. Belfield 2005). Ebenso finden sich empirische Hinweise darauf, dass der Unterricht in kleinen Klassen in den ersten Primarschuljahren einen nachhaltigen positiven Effekt auf Schulleistungen und Sekundarschulerfolg hat (vgl. Finn/Gerber/Boyd-Zaharias 2005). Positive Wirkungserwartungen werden in Deutschland mit dem Vorschlag verbunden, angesichts der begrenzten Arbeitsmarkteffekte der von der Bundesagentur für Arbeit geförderten beruflichen Weiterbildungsmaßnahmen Teile der dafür eingesetzten Mittel zur frühzeitigen schulischen Förderung von Risikogruppen einzusetzen.

Die Möglichkeiten einer intrasektoralen Mittelumerschichtung sollten nicht überbewertet werden. Zum einen fehlt es bislang an empirisch gesicherten Informationen über bereichsübergreifende Wirkungszusammenhänge und ihre Kostenimplikationen – solche Informationen könnte ein um Kostenanalysen erweitertes Bildungspanel liefern. Zum anderen sind zumindest kurz-

und mittelfristig solchen Mittelumschichtungen aufgrund der eingeschränkten Ausgabenflexibilität (der längerfristigen Ausgabenbindung) enge Grenzen gesetzt (Weishaupt/Weiß 1988).

### 3 Verstärkte Nutzung privater Finanzierungsquellen

Die verstärkte Inanspruchnahme privater Quellen wird die Finanzsituation nur verbessern, wenn die Mittel zusätzlich zur Verfügung stehen und nicht zu entsprechenden Kürzungen in den öffentlichen Haushalten genutzt werden. Die einzelnen Bildungsbereiche bieten für eine Aufstockung der Budgets durch private Mittel unterschiedliche Voraussetzungen. Im *Elementarbereich* geht die Tendenz derzeit eher dahin, die – nicht unerhebliche – private Kostenbeteiligung aus familien- und sozialpolitischen Gründen zu reduzieren. Anders stellt sich die Situation im *Hochschulsektor* dar: Hier folgt Deutschland mit der begonnenen Einführung von Studiengebühren der Mehrzahl der anderen OECD-Staaten. Sie sollen den Hochschulen als zusätzliche, im Wettbewerb um Studierende zu akquirierende Mittel zu Verbesserung der Qualität ihrer Lehre zur Verfügung stehen.

Was den *Schulbereich* betrifft, so sind die Möglichkeiten, über eine stärkere private Finanzierungsbeteiligung die Mittelausstattung nachhaltig zu verbessern, außerordentlich begrenzt. Keine Realisierungschance dürfte der Vorschlag des Sachverständigenrates Bildung (1998) haben, in der Sekundarstufe II Schulgeld einzuführen. Eine obligatorische private Finanzierungsbeteiligung ist im Wesentlichen auf die Einschränkung der Lernmittelfreiheit und die Erhebung von Gebühren bei der Nutzung zusätzlicher schulischer Dienstleistungen beschränkt. Bundesweit lässt sich ein Trend ausmachen, die Lernmittelfreiheit einzuschränken. Dies hat zu einer strukturellen Verschiebung der Finanzierungsanteile bei insgesamt schrumpfenden Lernmitteltats geführt (vgl. Avenarius et al. 2003, 124).

Auch die Möglichkeiten, über eine freiwillige private Finanzierungsbeteiligung die Mittelausstattung im Schulbereich nachhaltig zu verbessern, sind eher zurückhaltend einzuschätzen. Die für den allgemein bildenden Schulbereich vorliegenden Zahlen aus einer bundesweiten Stichprobenuntersuchung (Schmidt/Weishaupt/Weiß 2003) legen den Schluss nahe, dass – bezogen auf das gesamte Budget in diesem Bereich – „Drittmittel“ eine Quantité négligeable darstellen. In den disponiblen Budgets der einzelnen Schule können diese Mittel allerdings einen gewichtigen Anteil ausmachen. Den Angaben der Schulleitungen zufolge kann davon ausgegangen werden, dass im Schuljahr 1998/99 etwa die Hälfte der allgemein bildenden Schulen in Deutschland über keinerlei nennenswerte Drittmiteleinahmen verfügte. Im Durchschnitt standen den Schulen 7.012 DM (11,50 DM je Schüler) an zusätzlichen Mitteln (Mittel aus Fördervereinen, Elternspenden, Spenden- und Sponsorengelder privater Organisationen, geldwerte Sachleistungen) zur Verfügung, die vor allem zur Verbesserung der Ausstattung mit Lehr- und Lernmitteln sowie für außerunterrichtliche Aktivitäten verwendet wurden. Hinter diesem Betrag verbergen sich mit einer Spannweite von 260.000 DM allerdings beträchtliche Unterschiede zwischen den Schulen. Die darin sichtbar werdenden Chancenungleichheiten bei der Mobilisierung privater Ressourcen begründen in einem unter staatlicher Verantwortung stehenden Schulwesen Korrekturmaßnahmen (z. B. die Einrichtung regionaler Ausgleichsfonds), die dafür sorgen, dass die Disparitäten in den finanziellen Rahmenbedingungen der Schulen auf einen tolerierbaren Streubereich begrenzt werden.

## 4 Ausschöpfung von Effizienzreserven

### 4.1 Faktorielle Mittelverwendung

Der Ausschöpfung von Effizienzreserven wird in der aktuellen Finanzierungsdiskussion im Bildungsbereich relativ wenig Aufmerksamkeit geschenkt, obwohl es sich dabei um eine aussichtsreiche Strategie zur Entschärfung der Engpasssituation handelt, steht dahinter doch die ökonomische Einsicht, „that the closest one ever comes to getting something for nothing is by increasing efficiency (Harrold 1982, 58). Unter diesem Aspekt ist vor allem die *faktorielle Mittelverwendung* genauer in den Blick zu nehmen: die Verwendung der Mittel für die im schulischen Bildungsprozess eingesetzten Ressourcen. Innerhalb der bildungsökonomischen Forschung ist der Zusammenhang zwischen Bildungsressourcen und Bildungsqualität, meist indikatisiert durch Fachleistungen, Gegenstand einer Vielzahl empirischer Untersuchungen gewesen, die als „Produktivitätsstudien“ in die Literatur Eingang gefunden haben (vgl. z.B. Weiß/Timmermann 2004). Die Variablen-Beziehungen werden darin als Input-Output-Relationen einer „Bildungsproduktionsfunktion“ modelliert. Die für die ressourcenbezogenen Variablen (z.B. Ausgaben je Schüler, Lehrergehalt, Berufserfahrung der Lehrer, Schüler-Lehrer-Relation, Klassengröße, Sachmittelausstattung) vorgenommenen Parameterschätzungen sind ein Näherungswert ihrer jeweiligen „Grenzproduktivität“.

Die Ergebnisse von Produktivitätsstudien sind in verschiedenen Forschungsübersichten zusammengefasst worden. Besondere Publizität haben die mehrfach aktualisierten Auswertungen US-amerikanischer Studien durch Hanushek (1997) erlangt. Danach lässt sich kein enger und konsistenter Zusammenhang zwischen Schulressourcen und Schülerleistungen nachweisen. Dieses für die USA gezogene Forschungsresümee korrespondiert weitgehend mit den Ergebnissen einer Auswertung von Daten aus über 40 Teilnehmerländern der TIMS-Studie (vgl. Hanushek/Luque 2003). Real vorfindbaren Unterschieden in der Ausstattung mit finanziellen, personellen und materiellen Ressourcen kommt offenbar nur geringe Bedeutung als Varianzquelle von Schülerleistungen zu<sup>3</sup>.

Dieses – die herrschende Meinung innerhalb der Bildungsökonomie repräsentierende – Resümee ist in verschiedener Hinsicht zu differenzieren. Im Blick auf die hier interessierende Frage nach Möglichkeiten effizienzverbessernder Mittelumschichtungen ist erstens zu beachten, dass die einzelnen Ressourcen unterschiedliche Grenzproduktivitäten aufweisen, d.h. unterschiedlich leistungswirksam sind. Zweitens zeigen sich differenzielle Effekte für verschiedene Fächer und Schülergruppen. Drittens nähren Befunde methodisch anspruchsvollerer Untersuchungen zur Ressourcenwirksamkeit zum Teil Zweifel am Aussagegehalt der meist als Querschnittsanalysen angelegten Produktivitätsstudien (vgl. zu weiteren Kritikpunkten Weiß/Timmermann 2004).

Budgetneutrale Qualitätsverbesserungen lassen sich erreichen, wenn die Mittel zu den Ressourcen umgeschichtet werden, die eine höhere Grenzproduktivität je Geldeinheit (ein günstigeres Wirksamkeits-Kosten-Verhältnis) aufweisen. Im bildungsökonomischen Schrifttum finden sich nur wenige Studien, die empirische Informationen zur Produktivität einzelner Ressourcen mit Kosteninformationen zusammenführen, um zu Effizienzaussagen zu gelangen (z.B. Wolfe 1977; Pritchett/Filmer 1999; Levacic et al. 2005). Diese Studien verweisen auf ungenutzte Potenziale, durch Mittelum-

<sup>3</sup> Das Bild, das Querschnittsanalysen nationaler und internationaler Datensätze zum Einfluss von Schulressourcen auf Fachleistungen vermitteln, korrespondiert mit den Betrachtungen der längerfristigen Entwicklung des Zusammenhangs zwischen Ressourceneinsatz und Performanz von Schulsystemen in einer ganzen Reihe von Ländern. So lässt sich z.B. für die USA zeigen (Hanushek 2004), dass sich die preisbereinigten Ausgaben je Schüler zwischen 1960 und 2000 mehr als verdreifacht haben – bei bestenfalls konstant gebliebenen Testleistungen. Zu einem ähnlichen Ergebnis gelangen für OECD-Staaten Gundlach/Wößmann/Gmelin (2000).

schichtungen höhere Leistungsniveaus zu erreichen<sup>4</sup>. Ineffizienz bei der faktoriellen Mittelverwendung wird meist mit der bildungsbereichsspezifischen Technologie-Vagheit begründet (z.B. Murnane/Nelson 1984; dazu auch unten, Abschnitt 3.2). Pritchett/Filmer (1999) sehen dagegen den Hauptgrund dafür in einer interessengeleiteten Allokationspolitik, einer Mittelverwendung, die sich stärker an den Wohlfahrtszielen der Lehrerschaft als am Ziel der Verbesserung der Schuleffektivität orientiert. Sie begründen dies mit den Befunden einer unter Effizienzaspekten vorgenommenen Sekundäranalyse der Daten von Produktivitätsuntersuchungen in Entwicklungsländern. Danach finden die Mittel vorrangig für Inputs Verwendung, die direkt den Nutzen der Lehrer erhöhen (Anhebung der Lehrergehälter, Verkleinerung der Klassen), die aber gegenüber den von ihnen wenig präferierten Inputs (z.B. Ausgaben für Lehr- und Lernmittel) eine zehn- bis hundertmal niedrigere Grenzproduktivität je Dollar aufweisen. Auf den die Effizienz beeinträchtigenden Einfluss von Gewerkschaften auf die Allokationspolitik im Schulbereich verweist eine US-amerikanische Studie (Hoxby 1994).

Verschiedene Produktivitätsstudien lassen auf die Existenz *differenzieller Ressourceneffekte* schließen. Auf die Leistungen in „typischen“ Schulfächern wie Mathematik und Naturwissenschaften ist eher ein Einfluss von Ressourcen zu erwarten als auf Leistungen in Fächern, die stärker vom Elternhaus geprägt sind (z.B. Levacic et al. 2005). Hinweise finden sich auch auf eine Effektabhängigkeit von Schülermerkmalen. So konnte in amerikanischen Studien verschiedentlich festgestellt werden, dass vor allem leistungsschwächere Schülerinnen und Schüler sowie Kinder aus sozial und ethnisch benachteiligten Elternhäusern von kleinen Grundschulklassen profitieren (z.B. Finn/Gerber/Boyd-Zaharias 2005). Levacic et al. (2005) gelangen zu dem Ergebnis, dass sich der Einsatz zusätzlicher Mittel stärker auf die Leistungen von Schülerinnen und Schülern mit durchschnittlichen Fähigkeiten auswirkt als auf die Leistungen der oberen oder unteren zwanzig Prozent der Fähigkeitsverteilung. Eine an differenziellen Grenzproduktivitäten orientierte zielgruppenspezifische Ressourcenzuweisung würde nicht nur die Effizienz der Ressourcenallokation verbessern, sondern durch ihre Varianzreduzierende Wirkung zugleich zu größerer ergebnisbezogener Chancengleichheit beitragen. In Programmen „positiver Diskriminierung“ mit indikatorengesteuerter Ressourcenzuweisung, wie sie vielerorts existieren, haben solche Überlegungen Eingang gefunden (vgl. z.B. Ross/Levacic 1999).

Herkömmliche Produktivitätsstudien sehen sich mit einer ganzen Reihe von konzeptionellen und methodischen Problemen konfrontiert, die ihre Eignung für faktorielle Allokationsentscheidungen einschränken. Zu nennen sind u.a. die Beschränkung auf Fachleistungen (dazu ausführlich 3.2), das querschnittliche Untersuchungsdesign, die theoretisch nicht zu begründende Konzeptualisierung schulischer Ressourcen als direkte Wirkfaktoren und ungelöste Endogenitätsprobleme (vgl. ausführlich Weiß/Timmermann 2004, 252 ff.). Von Kritikern der Studien wird denn auch auf abweichende Befunde methodisch elaborierterer Untersuchungen verwiesen. Besondere Erwähnung verdient die in den USA im Bundesstaat Tennessee durchgeführte Experimentalstudie STAR zur Leistungswirksamkeit unterschiedlich großer Klassen im Primarbereich (Finn/Voelkl 1992, Class Size 1999). Im Unterschied zu konventionellen Schätzungen von Klassengrößen-Effekten in Bildungsproduktionsfunktionen wird in dieser Studie ein bedeutsamer, nachhaltig wirksamer Einfluss kleiner Klassen auf Schülerleistungen in Mathematik und Lesen festgestellt, der sich auch auf den späteren Sekundarschulabschluss positiv auswirkt (Finn/Gerber/Boyd-Zaharias 2005). Stellt man dem jedoch den beträchtlichen Aufwand einer flächendeckenden Verkleinerung der Klassen auf ein effektrelevantes Niveau (im Projekt STAR 13–17 Schüler) gegenüber (z.B. Brewer et al. 1999), sind erhebliche Zweifel an der Effizienz dieser Maßnahme angebracht.

<sup>4</sup> Eine englische Untersuchung (Levacic et al. 2005) zum Einfluss von Ressourcen auf die Leistungsentwicklung von Schülerinnen und Schülern empfiehlt z.B., mehr Mittel für Bücher als für Computer und zusätzliches Personal auszugeben.

## 4.2 Alternative Finanzierungskonzepte

Hohes qualitäts- und effizienzverbesserndes Potenzial wird Finanzierungskonzepten zugeschrieben, die als konstitutives Element neuer Steuerungssysteme Dezentralisierung der Ressourcenverantwortung und eine partiell schülergesteuerte Mittelzuweisung vorsehen. Im Rahmen institutioneller Finanzierung werden den Schulen mehr oder weniger umfassende Globalbudgets (formelgebunden) zugewiesen, ein unterschiedlich hoher Anteil der Mittel auf der Grundlage der gemeldeten Schülerzahlen (Pro-Kopf-Zuweisung), um Wettbewerb zwischen Schulen zu induzieren<sup>5</sup>. Den Wechsel zu subjektbezogener Finanzierung vollziehen mit derselben Intention propagierte Konzepte, bei denen die Bildungsnachfrager durch Aushändigen eines Bildungsgutscheins („vinkulierten Namenspapiers“) mit Kaufkraft ausgestattet werden<sup>6</sup>. Den Erwartungen nach sorgen die Anreizstrukturen dieser Finanzierungskonzepte für eine effizientere Mittelverwendung durch die schulischen Akteure. Die folgenden Ausführungen setzen sich kritisch mit dieser These auseinander. Dabei soll zunächst das der Operationalisierung des Effizienzkonzepts üblicherweise zugrunde liegende Verständnis von Schulqualität, dem auch die bisherigen Darstellungen folgten, problematisiert werden.

### *Eingführung auf Fachleistungen*

Jedes Verständnis von Schulqualität hängt von pädagogischen und bildungspolitischen Wertgesichtspunkten ab (vgl. Terhart 2000; Winch 1996). Den empirischen Bestimmungsversuchen von Schulqualität gehen also mehr oder weniger explizite Selektionsentscheidungen darüber voraus, was überhaupt als Ausdruck von Schulqualität erachtet wird, welche Indikatoren für deren empirische Bestimmung in Betracht kommen, welche davon man zum Zwecke der Überprüfung heranzieht und welche man (vorerst) nicht berücksichtigt und wie man diese Indikatoren gewichtet. Bei Versuchen der Operationalisierung von Schulqualität wird stets betont, dass man nicht beansprucht, das gesamte Leistungsspektrum schulischer Bildung in den Blick zu nehmen. Auch bei PISA wird bekanntlich nicht beansprucht, „den Horizont moderner Allgemeinbildung zu vermessen, oder auch nur die Umrisse eines internationalen Kerncurriculums nachzuzeichnen“ (Baumert et al. 2001, 21). Stattdessen konzentriert man sich auf wenige Basiskompetenzen, denen mit guten Gründen eine besondere Relevanz für schulische und außerschulische Bildungsbiographien beigemessen wird.

Leitender Gesichtspunkt der Problematisierung des Zusammenhangs von Bildungsfinanzierung und Schulqualität ist nun die Vermutung, dass sich die neuen Steuerungsinstrumente und die damit verknüpften Finanzierungsmodelle gegenüber dem Objekt der Steuerung nicht neutral verhalten, d.h. es wird nicht einfach das, was bislang und herkömmlicherweise unter schulischer Bildung verstanden wurde, effizienter, sondern *unter dem Einfluss dieser Steuerungsinstrumente verändert sich zugleich das Leistungsspektrum von Schulen sowie das Verständnis schulischer Bildung und ihrer Aufgaben*<sup>7</sup>. Wenn unter dem Einfluss neuer Finanzierungsmodelle Testleistungen in ausgewählten Leistungsdomänen steigen, muss also sehr genau untersucht werden, ob

<sup>5</sup> In England z.B. werden den Schulen von den Local Educational Authorities 80 % der Mittel der Globalbudgets auf der Basis der rekrutierten Schüler zugewiesen, 20 % auf der Basis schulspezifischer Faktoren: curricularen Besonderheiten, schülerbezogenem Sonderbedarf und schulstandortbedingtem Sonderbedarf (z.B. Zuschlag für kleine Schulen).

<sup>6</sup> Hierzu ausführlich *Dohmen*, Empirische Erfahrungen mit Bildungsgutscheinen: Was können wir für die deutsche Diskussion lernen?, in diesem Heft, S. 37.

<sup>7</sup> Darauf verweist eine kürzlich erschienene Studie des Center on Education Policy zu den Wirkungen des von der Bush Administration eingeführten *No Child Left Behind Act*. Danach berichten 71 % der 299 befragten Schulbezirke, dass die stärkere Konzentration auf testrelevante Fächer zu Kürzungen bei anderen Fächern geführt habe (vgl. *Stark Rentner* 2006).

und inwiefern die gemessenen Leistungssteigerungen als reale Effizienzgewinne interpretiert werden können und welche Kosten diesen Gewinnen gegenüberstehen.

Eine auf die neuen Steuerungsinstrumente ausgerichtete Bildungsfinanzierung verbindet die in Tests gemessenen Leistungen der Akteure und Organisationen mit Belohnungen und Sanktionen. Die gemessenen Leistungen können somit nicht nur rechtliche und administrative, sondern auch finanzielle Konsequenzen für Akteure und Organisationen haben. Auf Seiten der Akteure denke man etwa an Zulagen für herausragende Unterrichtsleistungen der Lehrer oder an leistungsbezogene Geldprämien für Schüler, wie sie etwa das in Michigan eingeführte *Merit Award Program* oder in England die „*GCSE bonuses*“ vorsehen<sup>8</sup>. Zu entsprechenden Effekten auf Seiten der Organisationen führen Finanzierungsmodelle mit Pro-Kopf-Zuweisungen: Schulen, die bei der Rekrutierung von Schülern erfolgreicher sind, erhalten unmittelbare finanzielle Vorteile.

Eine an Output- und Wettbewerbssteuerung orientierte Bildungsfinanzierung führt also zusätzliche *High Stakes* ein, die Anreize zu zusätzlichen Leistungen geben sollen. Aus der empirischen Bildungsforschung zu *High-Stakes Testing* ist freilich bekannt, dass entsprechende Reformstrategien problematische nicht-intendierte Nebenfolgen haben (vgl. Nichols/Glass/Berliner 2006). Politikempfehlungen von Seiten der Bildungsökonomie bezüglich neuer Finanzierungsmodelle können sich also nicht bloß darauf konzentrieren, welche Effekte diese Modelle auf einschlägige Qualitätsindikatoren haben; sie müssen sich vielmehr sehr sorgfältig damit auseinandersetzen, inwiefern die Einführung von *High Stakes* das Verhalten der Akteure und Organisationen im pädagogischen Feld insgesamt beeinflusst. Diese kritische Forschungsperspektive soll im Folgenden an zwei mit neuen Finanzierungsmodellen immer wieder verknüpften Erwartungen verdeutlicht werden: der Steigerung von Allokations- und Produktionseffizienz.

### *Allokative Effizienz*

Von Wettbewerbssteuerung und einer hiermit verknüpften Nachfragefinanzierung verspricht man sich insbesondere eine Diversifizierung des Bildungsangebots entsprechend den unterschiedlichen Präferenzen von Schülern und Eltern, was über die damit verbundenen Wohlfahrtsgewinne die allokative Effizienz steigern. Gegen Wettbewerbssteuerung im allgemeinbildenden Schulsystem ist zunächst ein normativer Einwand erhoben worden: Eine obligatorische Grundbildung sowie dessen öffentliche Finanzierung kann nur dadurch gerechtfertigt werden, dass sie ihr Angebot gerade *nicht* zur freien Wahl individueller Präferenzen stellt (vgl. Winch 1996). Abgesehen von diesem normativen Einwand, demzufolge Schulbildung als „meritorisches Gut“ zu begreifen ist, zeigen empirische Untersuchungen, dass Wettbewerb im allgemeinbildenden Schulsystem keineswegs zwangsläufig zur gewünschten Diversifizierung des Angebots führt (vgl. Lubienski 2003, 2006; LeTendre et al. 2001; Whitty 1997).

Aufschlussreich hierbei sind Studien über *Charter Schools*, bei denen man auf Grund ihrer Arbeitsbedingungen einen hohen Grad der Diversifizierung des Angebots erwarten würde. *Charter Schools* sind von direkten staatlichen Regularien weitgehend entlastet und stehen in einem freien Wettbewerb um Schüler, deren Schulbesuch öffentlich durch Pro-Kopf-Zuweisungen fi-

<sup>8</sup> Nach dem *Merit Award Program* erhalten Sekundarschüler ein einjähriges Stipendium in Höhe von 2.500 Dollar für den späteren Collegebesuch, wenn sie vorgegebene Leistungsstandards in vier Kompetenzbereichen (Mathematik, Naturwissenschaften, Lesen und schriftliche Sprachproduktion) erreichen oder überschreiten (vgl. Bishop 2002). Ähnliche Anreizstrukturen wurden in England für Schulen mit unterdurchschnittlichen Prüfungsergebnissen im Rahmen des Programms *New Deal for Communities* geschaffen. In Schulen, die an dem Programm teilnehmen, erhalten Schüler, die in fünf GCSE-Prüfungsfächern Noten zwischen A\* und C erreichen, 150–200 Pfund. Der Erfolg des Programms lässt sich daran ablesen, dass von den Schulen Summen zwischen 20.000 und fast 40.000 Pfund ausgezahlt werden (<http://www.telegraph.co.uk/news/main.jhtml;jsessionid=>

nanziert wird. Im Gegenzug zu ihrem größeren Gestaltungsspielraum steht eine strenge Rechenschaftspflicht über erzielte Leistungen.

Eine Sekundäranalyse von 190 US-amerikanischen Studien zu *Charter Schools* konnte die Hoffnung auf Innovation und Diversifizierung nicht bestätigen (vgl. Lubiencki 2003). Es lassen sich zwar einige Neuerungen auf administrativer Ebene feststellen (zum Beispiel 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 größeren Gestaltungsspielraums häufig durch eher traditionelle Unterrichtsmethoden und Lehrpläne aus. Generell zeigen international vergleichende Studien (LeTendre et al. 2001), dass die Routinen unterrichtlicher Interaktion (*classroom level*) eine noch geringere Varianz aufweisen als die Organisationsform von Schule (*administrative level*). Damit stellt sich die Frage, in welchem Maße neue Steuerungsinstrumente und damit verknüpfte Finanzierungsmodelle als distale Kontextfaktoren überhaupt bis auf die Ebene des Unterrichts durchgreifen.

Einige Bildungsforscher vermuten nun, dass es nicht trotz, sondern gerade infolge von Wettbewerbssteuerung zu der oft diagnostizierten Homogenisierung zugunsten eher traditioneller curricularer und unterrichtlicher Modelle kommt. Ein Erklärungsansatz für dieses erwartungswidrige Ergebnis stammt aus der neo-institutionalistischen Organisationsforschung (vgl. Schäfers 2002). In einem vielbeachteten Beitrag haben Paul DiMaggio und Walter Powell (1983/2000) auf Angleichungsprozesse hingewiesen, die sie als *institutionelle Isomorphie* bezeichnen. Organisationen operieren gemeinsam mit Mitbewerbern auf organisationalen Feldern, in denen es stets auch um die Beschaffung von Legitimitätsressourcen geht. Dabei kommt es zu Angleichungsprozessen, die im Einzelnen von drei Faktoren bestimmt werden: (a) Zwang durch staatliche und rechtliche Regulierungen, die das gesamte organisationale Feld bestimmen, (b) normativer Druck durch die Standards der Profession, (c) Mimesis, also die Nachahmung anderer vermeintlich erfolgreicher Organisationen und ihrer Problemlösungen, die ebenfalls im selben Feld operieren.

Gerade die Einführung von Wettbewerbselementen kann unter bestimmten Bedingungen Anreize für mimetischen Isomorphismus setzen. Eine treibende Kraft der Mimesis ist vor allem Unsicherheit. Im Einzelnen werden drei Gründe genannt: Unsicherheit entsteht zum Beispiel dann, wenn die Technologien einer Organisation nicht ausgereift, wenn ihre Ziele nicht eindeutig sind oder wenn die Umwelt symbolische Unsicherheit produziert. Auf der Grundlage dieser Analyse könnte man nun folgern, dass gegenwärtige Bildungsreformen zusätzliche Unsicherheiten ins Schulsystem hineinbringen, die vermutlich mimetischen Isomorphismus weiter begünstigen.

Generell besteht im pädagogischen Feld eine hohe Unsicherheit bezüglich der eigenen Technologien. Niklas Luhmann zufolge sind weder Individuen noch das Interaktionssystem Unterricht „Trivialmaschinen“, bei denen Input und Output, Aufwand und Ertrag in rationaler Weise kontrolliert und in Beziehung gesetzt werden könnte. Der Grund hierfür liegt in der unaufhebaren pädagogischen Differenz zwischen Lernen und Erziehen (vgl. Prange 2004), oder, systemtheoretisch gesprochen, in der Differenz zwischen der Operationsweise eines psychischen Systems (das Bewusstsein des Lernenden) und der Operationsweise eines sozialen Systems (das Interaktionssystem Unterricht). Für Luhmann ist das hieraus resultierende „Technologiedefizit“ auch Grund dafür, dass Organisationen im Erziehungssystem sich nicht durch ihr Produkt rechtfertigen lassen, sondern auf eine pauschale gesellschaftliche Anerkennung angewiesen sind (vgl. Luhmann 2002, 157). Die traditionelle Inputsteuerung und die damit verknüpfte Angebotsfinanzierung kann als ein Versuch verstanden werden, dem durch das Technologiedefizit mit-

bedingten Legitimitätsproblem der Erziehung durch pauschale Mittelzuweisung Rechnung zu tragen.

Genau diese pauschale gesellschaftliche Anerkennung scheint nun im Kontext gegenwärtiger Bildungsreform infrage gestellt zu werden. Man verlässt sich nicht mehr auf bloße Inputsteuerung, sondern macht die Zuweisung von materiellen und legitimatorischen Ressourcen davon abhängig, dass Ziele und Zielerreichung spezifiziert und für den außenstehenden Beobachter dokumentiert werden. Das Dilemma dabei ist, dass durch diese neue Form des ‘Controlling’ pädagogischer Organisationen das Technologiedefizit auf der Ebene pädagogischer Interaktion nicht behoben wird. Da die Systemebenen Organisation und Interaktion nur lose gekoppelt sind, führt eine veränderte Steuerung der Organisation Schule nicht unmittelbar zu den gewünschten Effekten auf der Ebene der Interaktion (vgl. Meyer/Rowan 1977/1991).

Outputsteuerung misst die Organisation folglich an Ergebnissen, für deren Zustandekommen keine eindeutigen Technologien zur Verfügung stehen. Die neuen Steuerungsmodelle erzeugen somit eine neue Form von Unsicherheit, die mimetischen Isomorphismus verstärkt. Man weiß nicht genau, wie man die gewünschten Ergebnisse hervorbringt, also ahmt man denjenigen nach, der die gewünschten Ergebnisse abliefert, unbeschadet der Tatsache, dass vermutlich auch dieser nicht genau weiß, wie er das macht. Unter Umständen kann diese Strategie sogar erfolgreich sein, ohne dass die Organisation ein Wissen über die Bedingungen ihres Erfolgs generiert hätte. Bisweilen genügt es auch, andernorts verbreitete Rationalitätsmythen in die Formalstruktur der Organisation aufzunehmen, während das Kerngeschäft des Unterricht weiter nach ‚bewährten‘ Routinen verfährt. Angesichts technologiebedingter Handlungsunsicherheit ist es im Übrigen allemal Erfolg versprechender, *High Stakes* über Klientenselektion zu sichern<sup>9</sup>.

Die Erwartung, Wettbewerbssteuerung und eine damit verknüpfte Nachfragefinanzierung führe zu einer Diversifizierung des Angebots und erhöhe damit die alloкатive Effizienz, lässt sich generell nicht bestätigen. Am Beispiel der *Charter Schools* und anderer Erfahrungen mit Wettbewerbssteuerung im Bildungssystem (vgl. Lubienski 2006) zeigt sich vielmehr ein auch in anderem Zusammenhang (vgl. Weiß 2002) beobachteter Sachverhalt: Institutionelle Akteure nutzen größere Gestaltungsspielräume nicht notwendigerweise für die Erzielung von Effizienzgewinnen. Unbestritten bleibt allerdings, dass allein die im Kontext von Wettbewerbssteuerung sich eröffneten Wahlmöglichkeiten für die Nachfrageseite ein Wohlfahrtsgewinn bedeuten kann. So lässt sich zeigen, dass die Zufriedenheit über Schulwahlmöglichkeiten zunimmt, auch wenn von der Wahlmöglichkeit gar kein Gebrauch gemacht wird bzw. das, was zu Wahl steht, sich wenig unterscheidet (vgl. Bellmann/Waldow 2006). Unbestritten ist aber auch, dass die Wohlfahrtserträge nicht sozialschichtneutral verteilt sind. Gewinner sind vor allem die Nachfrager, denen aufgrund ihrer besseren Ausstattung mit ökonomischem, kulturellem und sozialem Kapital anbieterseitig eine besondere Wertschätzung entgegengebracht wird und deren Präferenzen deshalb stärkere Berücksichtigung finden (vgl. Woods/Bagley/Glatler 1998). Wenn dies aber über Exklusion und Creaming dazu führt, dass eine Verbesserung der Wohlfahrtsposition bei bestimmten Individuen mit einer Verschlechterung der Wohlfahrtsposition bei anderen Individuen einhergeht, verbessert sich dem Pareto-Kriterium zufolge die gesellschaftliche Wohlfahrt nicht.

<sup>9</sup> In einer kürzlich erschienen englischen Studie (vgl. Slater/Stewart 2006) konnte gezeigt werden, dass Sekundarschulen vor allem über Eingangsselektivität und den Ausschluss schwieriger Schüler ihre Ergebnisse in den Abschlussprüfungen (GCSE) verbesserten.

### Produktionseffizienz

Im Kontext von *High-Stakes Testing* sind immer wieder überraschende Anstiege von Testleistungen festgestellt worden (vgl. Haney 2000). Dies lässt zunächst solche Finanzierungsmodelle attraktiv erscheinen, die versuchen, die Effekte von *High Stakes* durch finanzielle Anreizsysteme zu verstärken. Bevor entsprechende Politikempfehlungen gegeben werden, muss allerdings genauer untersucht werden, inwiefern steigende Testleistungen tatsächlich Ausdruck einer gesteigerten pädagogischen Produktivität und besserer Schulleistungen sind. Daniel Koretz (2002, 2005) zufolge ist es ein immer wieder beobachtbares Phänomen, dass Testergebnisse nach Einführung eines neuen sanktionsbewehrten Evaluationsverfahrens zunächst abfallen, dann deutlich ansteigen, um sich nach wenigen Jahren auf einem ähnlich hohen Niveau zu stabilisieren, auf dem auch das vorher zum Einsatz gekommene Evaluationsverfahren angelangt war. Hinter diesem Phänomen können, Koretz zufolge, sehr unterschiedliche Verhaltensweisen der Akteure im pädagogischen Feld stecken, worunter sowohl erwünschte als auch unerwünschte Reaktionsformen fallen.

Eindeutig erwünscht wäre es beispielsweise, wenn Lehrer bessere Arbeit leisten, indem sie etwa auf effektivere Unterrichtsmethoden zurückgreifen. Eindeutig unerwünscht wären dagegen alle Formen des Betrugs, etwa indem vorab an Testitems geübt wird, Hilfestellungen während des Tests gegeben werden oder falsche Antworten korrigiert werden<sup>10</sup>. Zwischen diesen beiden Extremen gibt es eine Grauzone von Vorbereitungen auf den Test, die weniger eindeutig zu bewerten sind. „Teaching to the test“ kann dabei keineswegs pauschal als unerwünscht bezeichnet werden. Es ist vielmehr ein unscharfer Sammelbegriff für unterschiedliche Verhaltensweisen, mit denen Akteure auf *High-Stakes Testing* reagieren und die im Ergebnis zu steigenden Testleistungen führen. In diese Grauzone fallen z. B. die Umverteilung von Unterrichtszeit innerhalb und zwischen Leistungsdomänen (*reallocation*), die inhaltliche Ausrichtung des Unterrichts auf den Test (*alignment*) sowie die Ausrichtung des Unterrichts am Format von Testaufgaben (*coaching*). Alle drei Varianten können im Ergebnis zu „score inflation“ führen. Das Resümee lautet daher: „Absent additional information, one cannot tell from aggregate data what mix of these responses occurred, and one therefore cannot judge how meaningful the gains are“ (Koretz 2002, 764). Als Bildungsforscher kann sich Koretz daher nur wundern, wenn Zugewinne in Testleistungen von Politik und Öffentlichkeit als unzweideutige Indikatoren für gestiegene Schülerleistungen und höhere Produktivität von Schulen interpretiert werden.

Die Konzentration des Tests auf wichtige und curricular valide Unterrichtsinhalte ist dabei noch kein hinreichender Schutz vor der Inflation von Testleistungen, denn es muss zugleich gefragt werden, was im Zuge der Reallokation von Unterrichtszeit *weniger* Gewicht erhält: „given the time constraints confronting teachers, [...] an increase in emphasis on the tested elements would lead to a decrease in emphasis on other, untested elements. If some of these untested elements are important for the intended inferences, then the increase in scores could mask either no change or an actual decrease in mastery of the domain the test is supposed to represent.“ (Koretz 2005, 13) Da aber sowohl der Bestimmung von domänenspezifischen Testitems als auch der Bestimmung von Indikatoren für Schulqualität insgesamt notgedrungen Auswahlentscheidungen zu Grunde liegen, stellt die Fokussierung des Unterrichts auf ausgewählte Qualitätsindikatoren eine zumindest ambivalente Reaktionsform der Akteure dar. Dabei zeigt sich auch, dass Validität keine Eigenschaft darstellt, die einem Test als solchem zukommt; sie kann sich vielmehr im Kontext der Reaktionen der Akteure im Umgang mit dem Test so verändern,

<sup>10</sup> Auffallend ist, dass sich Meldungen über solche Formen der Einflussnahme mit der Einführung von *High-Stakes Testing* häufen (vgl. Popham 2006). Auch in Deutschland werden im Zusammenhang mit dem Einsatz des Vera-Tests in den Grundschulen einiger Länder Betrugsvorwürfe laut (vgl. „Der Spiegel“ vom 3.7.2006, S. 52/53).

dass ein anfänglich valider Test allmählich seine Aussagekraft verlieren und als Indikator für tatsächliche Leistungssteigerungen unbrauchbar werden kann.

Die verstärkte Orientierung an quantifizierbaren Indikatoren für Schulqualität sowie die Kopplung von Evaluationsergebnissen an *High Stakes* birgt Risiken, die inzwischen auch empirisch gut belegt sind. Nichols und Berliner (2005) sprechen im Anschluss an Donald T. Campbell von der Gefahr einer Korruption sowohl der Indikatoren als auch der Akteure, deren Leistung mit diesen Indikatoren gemessen werden sollen. „The more any quantitative social indicator is used for social decision-making, the more subject it will be to distort and corrupt the social processes it is intended to monitor.“<sup>11</sup>

Die hiermit angedeuteten Risiken einer an quantitativen Indikatoren orientierten Qualitätsentwicklung verdient auch in der Bildungsökonomie eine größere Aufmerksamkeit. Hierfür sprechen verschiedene Gründe: Wenn Outputsteuerung durch *High-Stakes Testing* empirisch nachweisbare nicht-intendierte Nebenfolgen (wie „score inflation“ bzw. eine Korruption sowohl der Qualitätsindikatoren als auch der Akteure im pädagogischen Feld) *systematisch* erzeugt, dann wären diese Nebenfolgen als Kosten des neuen Steuerungssystems – zusätzlich zu den meist vernachlässigten Transaktionskosten (vgl. Weiß 2003)<sup>12</sup> – zu berücksichtigen. Wenn zur Ausschöpfung von Effizienzreserven *High Stakes* in Form von finanziellen Anreizsystemen für Akteure und Organisationen eingeführt werden, dann steigt das Risiko, bloß *scheinbare* Effizienzgewinne zu erzielen. Dem bildungspolitischen Rationalitätskalkül, durch Mittelumrichtungen das Leistungsniveau kostenneutral anzuheben, antwortet dann auf Seiten der Akteure im pädagogischen Feld ein vergleichbares Kalkül: Auch hier hat man es mit begrenzten, vor allem zeitlichen Ressourcen zu tun, so dass die Akteure versuchen werden, das Leistungsniveau (in den getesteten Domänen) durch Reallokation der ihnen zur Verfügung stehenden Ressourcen in einer Weise anzuheben, die aus ihrer Sicht kostenneutral ist (vgl. Fußnote 5). ‚Echte‘ Effizienzgewinne sind dagegen nur unter bestimmten Bedingungen wahrscheinlich, die im Rahmen neuer Steuerungsinstrumente nicht oder nur sehr unvollkommen realisiert sind. Eine unabdingbare Voraussetzung ist dabei ein normativer Bezugsrahmen für Evaluationen, dem ein möglichst breites und vielschichtiges Verständnis von Schulqualität zu Grunde liegt, das mit verschiedenen Indikatorenbündeln bestimmt wird, um den Anreiz zur Erzielung nur scheinbarer Effizienzgewinne entlang einiger weniger Leistungsparameter zu verringern.

Eine kritische Rückfrage an neue Steuerungsinstrumente und die damit verbundenen Finanzierungsmodelle wäre folglich auch, inwiefern sie systematische Anreize zur Privilegierung *bestimmter* Ertragsdimensionen von Bildung enthalten und somit einem eingeschränkten Verständnis von Schulqualität Vorschub leisten. Der Bildungsökonom Henry Levin (2000) macht immer wieder darauf aufmerksam, dass Bildungssysteme unterschiedliche Aufgaben gleichzeitig zu erfüllen haben: Neben der effizienten ‚Produktion‘ von Wissen und Können, sollen sie Eltern und Schülern Wahlmöglichkeiten eröffnen, zugleich Bildungschancen unabhängig von sozialer Herkunft gewähren und durch allgemeine Bildung eine gemeinsame Grundlage demokratischen Zusammenlebens erzeugen. Eine Reform zugunsten einer dieser Aufgabendimensionen geht freilich mitunter zu Lasten einer anderen Dimension. Es kann Zielkonflikte oder *trade-offs* geben, die in der Kosten-Nutzenbilanz des

<sup>11</sup> Campbell 1975, S. 35.

<sup>12</sup> Solche Transaktionskosten (Kosten der „Benutzung“ eines Steuerungssystems) entstehen etwa durch zusätzlichen Koordinations- und Kontrollaufwand sowie notwendige nachträgliche Korrekturen unerwünschter Wettbewerbsergebnisse in einem unter staatlicher Verantwortung stehenden Schulsystem. Ein Beispiel dafür sind die zahlreichen Agenturen zur Platzierung von Schülern („admission authorities“), die in England eingeführt wurden, um dem durch die Schulwahlfreiheit entstandenen Planungschaos Herr zu werden. Auf die erheblichen Kosten der Administration eines gutscheinbasierten Finanzierungssystems verweist eine für die USA vorgenommene Modellrechnung (Levin 1998).

neuen Steuerungsmodells berücksichtigt werden müssen. Eine Fokussierung auf bestimmte Ertragsdimensionen schulischer Bildung ist dabei nicht prinzipiell problematisch – öffentliche Schulbildung kann selbst nur einen eingeschränkten Beitrag zur Bildung in einem umfassenderen Sinne leisten –, aber die Fokussierung wäre transparent zu machen, innerhalb eines normativen Rahmens zu diskutieren und demokratisch legitimierten Entscheidungen zugänglich zu machen. Es genügt jedenfalls nicht, wenn sich eine solche Fokussierung als nicht-intendierte Nebenfolge neuer Steuerungssysteme gewissermaßen hinterrücks einstellt. Keineswegs lässt sich die Inkaufnahme nicht-intendierter Steuerungseffekte mit bedeutsamen intendierten Wirkungen begründen, fallen doch die dazu vorliegenden empirischen Befunde – etwa zum Einfluss erweiterter schulischer Dispositionskompetenz über Ressourcen und nachfrageorientierter Finanzierungskonzepte auf Testergebnisse in Basiskompetenzen – widersprüchlich aus (vgl. Weiß 2003; zur Wirkung von Bildungsgutscheinen Belfield 2001; Dohmen 2005<sup>13</sup>). Auch hinsichtlich eines engen Verständnisses von Schulqualität ist die Forschung bislang einen überzeugenden Nachweis des effizienzverbessernden Potenzials der im neuen Steuerungsregime vorgesehenen Finanzierungskonzepte schuldig geblieben.

## 5 Resümee

Hinsichtlich der in der Überschrift formulierten Frage lässt sich resümieren:

- Jenseits einer stets „subjektiv gefühlten“ Unterausstattung mit Ressourcen ist zu konstatieren: Die Beurteilung der Finanzsituation im Schulbereich fällt unterschiedlich aus, je nachdem, welcher Bewertungsmaßstab, welche Indikatoren und Benchmarks herangezogen werden. Gemessen an dem von Experten für Qualitätsverbesserungen für notwendig erachteten Ressourcenbedarf ist der Schulbereich unterfinanziert.
- Derzeit spricht wenig dafür, dass der zusätzliche Mittelbedarf des Schulwesens durch wachstumsinduzierte Mehreinnahmen oder seine Besserstellung im gesamtstaatlichen Mittelverteilungsprozess gedeckt werden kann. Bei relativ expansionsinflexiblen Bildungsbudgets rücken andere strategische Ansatzpunkte in den Blickpunkt: intrasektorale Mittelumschichtungen, die verstärkte Nutzung privater Finanzierungsquellen und die Ausschöpfung von Effizienzreserven.
- Auch wenn Plausibilität und zum Teil vorliegende empirische Befunde eine prioritäre Mittelverwendung für fokussierte Qualitätsverbesserungen in den unteren Bildungsstufen nahe legen, fehlt bislang eine tragfähige informationelle Grundlage für evidenzbasierte Mittelumschichtungen innerhalb des Bildungssystems. Erschwert wird eine solche Reallokation der Mittel durch die eingeschränkte Ausgabenflexibilität.
- Wenig Möglichkeiten bieten sich im Schulbereich, die Mittelausstattung durch die verstärkte Erschließung privater Finanzierungsquellen nachhaltig zu verbessern. Auch wenn im Rahmen strategischer Überlegungen zur Ressourcensicherung die Mobilisierung privater Mittel nicht mehr wegzudenken ist, besteht doch breiter Konsens darüber, dass dies in einem unter staatlicher Gesamtverantwortung stehenden Schulwesen kein konstitutives Element eines „neuen“ Finanzierungssystems sein kann. Evidente Disparitäten zwischen den Schulen beim privaten Mittelaufkommen begründen schon jetzt einen auf Chancenausgleich gerichteten politischen Handlungsbedarf.
- Die bildungsökonomische Forschung liefert hinreichende Evidenz dafür, dass Veränderungen des Ressourceneinsatzes innerhalb der real vorfindbaren Variationsbereiche keine sub-

<sup>13</sup> Vgl. insoweit auch *Dohmen*, Empirische Erfahrungen mit Bildungsgutscheinen: Was können wir für die deutsche Diskussion lernen?, in diesem Heft, S. 37.

stanzialen Qualitätseffekte, gemessen an Fachleistungen, nach sich ziehen. Zugleich finden sich aber Indizien für ungenutzte Potenziale budgetneutraler Qualitätsverbesserungen durch faktorielle Mittelumschichtung. Sollte die These zutreffen, dass sich darin nicht nur ein Technologiedefizit, sondern auch eine – dadurch begünstigte – interessen geleitete Allokationspolitik manifestieren, dürfte auch von dieser Strategie kaum ein Beitrag zur Entschärfung der finanziellen Engpasssituation im Schulbereich zu erwarten sein. Weniger politischem Widerstand ist indes eine durch die empirische Bildungsforschung gestützten indikatoren gesteuerten Mittelzuweisung im Sinne „positiver Diskriminierung“ ausgesetzt.

- Die kritische Diskussion und empirische Erforschung der Effekte von Finanzierungskonzepten des neuen Steuerungsparadigmas hat inzwischen eine Reihe von Ergebnissen hervorgebracht, die Anlass zu erheblichen Zweifeln am Effizienzverbesserungspotenzial dieser Konzepte geben. Abgesehen davon, dass bislang für die üblicherweise verwendeten Qualitätsindikatoren (Fachleistungen) ein eindeutiger empirischer Nachweis positiver Effizienzeffekte aussteht, sehen sich die neuen Finanzierungskonzepte mit dem Risiko einer problematischen Fehlsteuerung der Qualitätsentwicklung von Schulen behaftet, denn: Die ihnen inhärenten – das Verhalten der Akteure und Organisationen prägenden – Anreizstrukturen einer an Output und/oder Nachfrage geknüpften Mittelzuweisung erzeugen systematisch nicht-intendierte Nebenwirkungen. Bleiben diese unberücksichtigt, werden positive Ergebnisse bei evaluationsrelevanten Indikatoren (z.B. verbesserte Testleistungen) vorschnell als Steigerung der Produktionseffizienz interpretiert.
- Die an eine wettbewerbsorientierte Mittelzuweisung geknüpfte Erwartung größerer allokativer Effizienz auf Grund von individuellen Wohlfahrtsgewinnen eines stärker diversifizierten Bildungsangebots, das den Nachfragerpräferenzen besser entspricht, ist empirisch nicht zu sichern. Stattdessen zeigen sich Homogenisierungstendenzen, die ihre Ursache vor allem in technologiebedingter Handlungsunsicherheit und davon ausgehenden Anreizen für mimetischen Isomorphismus haben.

## Literatur

- Avenarius, H. et al., Bildungsbericht für Deutschland, Opladen 2003.
- Baumert, J. et al. (Hrsg.), PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich, Opladen 2001, S. 19–33.
- Belfield, C.R., International experience with demand-led financing: education vouchers in the USA, Great Britain/Chile/New York 2001.
- Belfield, C., The promise of early childhood education, New York 2005.
- Bellmann, J./Waldow, F., Bildungsgutscheine als Steuerungsinstrument. Egalitäre Erwartungen, segregative Effekte und das Beispiel Schweden, in: Ecarius, J./Wigger, J. (Hrsg.), Elitebildung – Bildungselite. Erziehungswissenschaftliche Diskussionen und Befunde über Bildung und soziale Ungleichheit, Opladen 2006, S. 188–205.
- Bishop, J.H., A prospective policy evaluation of the Michigan Merit Award Program, paper presented at the conference „Taking Account of Accountability“, Boston 2002.
- Bos, W. et al. (Hrsg.), Erste Ergebnisse aus IGLU, Münster u.a. 2003.
- Brewer, D. et al., Estimating the cost of national class size reductions under different policy alternatives, Educational Evaluation and Policy Analysis 21 (2), 1999, S. 179–192.
- Bundesministerium des Inneren, Dritter Versorgungsbericht der Bundesregierung. Berlin 2005.

- Campbell, D.T., Assessing the impact of planned social change, in: Lyons, G.M. (Hrsg.), *Social research and public policies. The Dartmouth/OECD Conference, Hanover/New Hampshire 1975*, S. 3–45.
- Class size: Issues and new findings, *Educational Evaluation and Policy Analysis (Special Issue)*, 21 (2), 1999.
- DiMaggio, P.J./Powell, W.W., Das ‘stahlharte Gehäuse’ neu betrachtet: Institutioneller Isomorphismus und kollektive Rationalität in organisationalen Feldern, in: Müller, H.-P./Sigmund, S. (Hrsg.), *Zeitgenössische amerikanische Soziologie, Opladen 1983/2000*, S. 147–174.
- Dohmen, D. *Theorie und Empirie von Bildungsgutscheinen: Was können wir für die deutsche Diskussion lernen? Köln 2005*.
- Finn, J./Voekl, K., Class size, in: Husen, T.& Postlethwaite, T.N. (Hrsg.), *International Encyclopedia of Education*, 2. Aufl. Oxford 1992, S. 770–775.
- Finn, J.D./Gerber, S.B/Boyd-Zaharias, J. (2005), Small classes in early grades, academic achievement, and graduating from high school, *Journal of Educational Psychology* 97 (2), 2005, S. 214–223.
- Gundlach, E./Wößmann, L./Gmelin, J., The decline of schooling productivity in OECD countries, *Economic Journal* 11, 2000, S. 135–147.
- Haney, W. The Texas miracle in education, *Education Policy Analysis Archives*, 41 (8), 2000. Verfügbar unter: <http://epaa.asu.edu/epaa/v8n41/> [19. Juni 2006].
- Hanushek, E.A., Assessing the effects of school resources on student performance: An update, *Educational Evaluation and Policy Analysis* 19 (2), 1997, S. 141–164.
- Hanushek, E.A./Luque, J.A., Efficiency and equity in schools around the world, *Economics of Education Review* 22, 2003, S. 481–502.
- Hanushek, E.A., *Distributional outcomes of the organization of U.S. schools: peers, school quality, and achievement, München 2004*.
- Harrold, R., *Economic thinking in education, Armidale 1982*.
- Hauptmeier, S., *Schuldenmonitor: Tragfähigkeit der öffentlichen Haushalte in Deutschland, Wachstums- und Konjunkturanalysen* 8 (4), 2005, S. 6–7.
- Homburg, S., *Nachhaltige Finanzpolitik für Niedersachsen, Hannover 2005*.
- Hoxby, C.M., How teachers’ unions affect education production, *Quarterly Journal of Economics* 111, 1994, S. 671–718.
- Klemm, K., *Bildungsausgaben in Deutschland: Status quo und Perspektiven, Bonn 2005*.
- Konsortium Bildungsberichterstattung (Hrsg.), *Bildung in Deutschland, Bielefeld 2006*.
- Koretz, D.M., Limitations in the use of achievement tests as measures of educators’ productivity, *The Journal of Human Resources* 37 (4), 2002, S. 752–777.
- Koretz, D., Alignment, high stakes, and the inflation of test scores, CSE Report 655, Center for the Study of Evaluation (CSE), UCLA, 2005. Verfügbar unter: <http://www.cse.ucla.edu/> [19.Juni 2006].
- Levacic, R. et al., *Estimating the relationship between school resources and pupil attainment at key stage 3, London 2005*.
- LeTendre, G.K./Baker, D.P./Akiba, M./Goesling, B./Wiseman, A., Teachers’ work: Institutional isomorphism and cultural variation in the U.S., Germany, and Japan, *Educational Researcher* 30 (6), 2001, S. 3–15.

- Levin, H.M., Educational vouchers: Effectiveness, choice, and costs, *Journal of Policy Analysis and Management* 17 (3), 1998, S. 373–392.
- Levin, H.M. (2000). Recent developments in the economics of education: educational vouchers, in: Weiß, M./Weishaupt, H. (Hrsg.), *Bildungsökonomie und Neue Steuerung*, Frankfurt/M. 2000, S. 97–114.
- Lubienski, C., Innovation in education markets: Theory and evidence on the impact of competition and choice in charter schools, *American Educational Research Journal* 40 (2), 2003, S. 395–443.
- Lubienski, C., School diversification in second-best education markets. International evidence and conflicting theories of change, *Educational Policy* 20 (2), 2006, S. 323–344.
- Luhmann, N., *Das Erziehungssystem der Gesellschaft*, Frankfurt a.M. 2002.
- Meyer, J.W., Rowan, B., Institutionalized organizations: Formal structure as myth and ceremony, in: Powell, W.W./DiMaggio, P.J. (Hrsg.), *The New Institutionalism in Organizational Analysis*, Chicago/London 1977/1991, S. 41–62.
- Murnane, R.J./Nelson, R.R., Production and innovation when techniques are tacit, *Journal of Economic Behavior and Organization* 5, 1984, S. 353–373.
- Nichols, S.L., Berliner, D.C., The inevitable corruption of indicators and educators through high-stakes testing. The great Lakes Center for Education Research & Practice, 2005, S. 1–187. Verfügbar unter [http://greatlakescenter.org/g\\_1\\_new\\_doc/EPSTL-0503-101-EPRU.pdf](http://greatlakescenter.org/g_1_new_doc/EPSTL-0503-101-EPRU.pdf)
- Nichols, S.L., Glass, G.V., Berliner, D.C., High-stakes testing and student achievement: Does accountability pressure increase student learning?, in: *Education Policy Analysis Archives* 14 (1), 2006, S. 1–172. Verfügbar unter <http://epaa.asu.edu/epaa/v14n1/> [19. Juni 2006].
- Popham, J.W. Educator cheating on No Child Left Behind Tests, *Education Week*, April 19, 2006, S. 32–33.
- Prange, K., Bildung: a paradigm regained? *European Educational Research Journal* 3 (2), 2004, S. 501–509.
- Pritchett, L., Filmer, D., What education production functions really show: A positive theory of education expenditures. *Economics of Education Review* 18, 1999, S. 223–239.
- Ross, K.N., Levacic, R. (Hrsg.), *Needs-based resource allocation in education via formula funding of schools*, Paris 1999.
- Sachverständigenrat: *Jahresgutachten 2004/05, Erfolge im Ausland – Herausforderungen im Inland*, Wiesbaden 2004.
- Sachverständigenrat Bildung (Hrsg.), *Für ein verändertes System der Bildungsfinanzierung*. Düsseldorf 1998.
- Schäfers, C., Der soziologische Neo-Institutionalismus. Eine organisationstheoretische Analyse- und Forschungsperspektive auf schulische Organisationen, *Zeitschrift für Pädagogik*, 48 (6), 2002, S. 835–855.
- Schmidt, C., Weishaupt, H., Weiß, M., Einzelschulische Bedingungen dezentraler Ressourcenverantwortung, *Recht der Jugend und des Bildungswesens*, 51 (1), 2003, S. 105–120.
- Slater, J., Stewart, W., Standards drive falters, say studies, *TES* vom 31.3., (2), 2006.
- Stark Rentner, D. et al. (2006). *From the capital to the classroom: Year 4 of the No Child Left Behind Act*. Washington: Center on Education Policy.
- Terhart, E., Qualität und Qualitätssicherung im Schulsystem. Hintergründe – Konzepte – Probleme, *Zeitschrift für Pädagogik*, 46 (6), 2000, S. 809–829.

- Vereinigung der Bayerischen Wirtschaft (Hrsg.), Bildung neu denken! Das Zukunftsprojekt, Wiesbaden 2003.
- Vereinigung der Bayerischen Wirtschaft (Hrsg.), Bildung neu denken! Das Finanzkonzept, Wiesbaden 2004.
- Weishaupt, H., Weiß, M., Bildungsbudget und interne Mittelallokation, Zeitschrift für Pädagogik, 34 (4). 1988, S. 535–553.
- Weiß, M., „Schools are special places“ – Grenzen der Neuen Steuerung im Schulbereich, Schule und Wirtschaft. Auf der Suche nach einem neuen Verhältnis. Münster 2002.
- Weiß, M., Kann das Schulwesen durch Wettbewerb genesen?, in: Döbert, H. et al. (Hrsg.), Bildung vor neuen Herausforderungen. Neuwied 2003, S. 111–119.
- Weiß, M., Timmermann, D., Bildungsökonomie und Schulstatistik, in: Helsper, W./Böhme, J. (Hrsg.), Handbuch der Schulforschung, Opladen 2004, S. 241–287.
- Whitty, G., Creating quasi-markets in education, in: Apple, M.W. (Hrsg.), Review of Research in Education, 1997, S. 3–47.
- Winch, C., Quality and education. Second Special Issue of the Journal of Philosophy of Education, Oxford 1996.
- Wolfe, B., A cost-effectiveness analysis of reductions in school expenditures: an application of an educational production function, Journal of Educational Finance, 1997, S. 407–418.
- Woods, P.A., Bagley, C., Glatter, R., School choice and competition: markets in the public interest? London 1998.
- Verf.: Prof. Dr. Manfred Weiß, Im Alten Grund 10, 65812 Bad Soden,*
- Prof. Dr. Johannes Bellmann, Institut für Erziehungswissenschaft, Pädagogische Hochschule Freiburg, Kunzenweg 21, 79117 Freiburg*