

Ausbildung oder Auswahl? Der Einfluss der Hochschule auf das Allgemeinwissen

Björn Frank, Rainer Voßkamp (Universität Kassel)

Zusammenfassung:

Der Studentenpisa-Test ermöglicht es, den Wissensstand von Studierenden zu beurteilen. Der Beitrag unternimmt mit Hilfe der Studentenpisa-Daten den Versuch, nicht nur den Wissensstand, sondern den Wissenszuwachs zu messen und zu erklären. Wächst das Allgemeinwissen der Studierenden an bestimmten Hochschulen während ihres Studiums stärker als an anderen Hochschulen? Und was kennzeichnet eine Hochschule, an der die Studierenden mehr hinzulernen? Der Beitrag konzentriert sich auf den möglichen Einfluss, den die Qualität von Forschung und Lehre einer Hochschule hat. Es zeigt sich, dass der Wissenszuwachs wenigstens für die Fächer Wirtschaft und Politik – nicht aber für Geschichte – mit bekannten Maßen für die Qualität der Lehre an Hochschulen zusammenhängt: je besser die Lehre, desto höher der Wissenszuwachs. Dieser Zusammenhang ist an Fachhochschulen stärker als an Universitäten. Die Qualität der Forschung an einer Universität hat dagegen keinen Einfluss auf den Wissenszuwachs ihrer Studierenden.

1. Einleitung

In diesem Beitrag werden wir mit statistischen Methoden ermitteln, welches die schönste deutsche Universität ist. Oups, vertippt. Aber das hätte uns ohnehin niemand abgekauft. Schönheit liegt ja im Auge des Betrachters. Oder doch nicht? Kann man nicht doch etwas messen? Sagen wir, die Größe der Grünflächen auf dem Campus, oder wie viel Tageslicht in die Hörsäle fällt? Ja, das ginge. Es würde bloß niemand behaupten, damit ließe sich wirklich die schönste Universität ermitteln.

Und so ähnlich ist es auch mit der „besten“ Universität. Wiederum liegt Qualität im Auge des Betrachters, und wiederum kann man zwar ein paar Dinge messen, darf dann aber nicht glauben, die Qualität damit wirklich objektiv erfasst zu haben. Daher müssen wir vorsichtig sein, wenn wir dennoch dieses Spiel mitspielen: Wir versuchen, einen bestimmten Aspekt der Qualität von Hochschulen zu messen, und zwar mit Hilfe der Studentenpisa-Daten. Unser Ziel ist es, den Wissensunterschied zwischen Studienanfängern und Studierenden höherer Semester zu messen und zu erklären. Das Studium an sich wird nicht die einzige Ursache für diese Wissensunterschiede sein, aber wenn an einer Hochschule der Wissenszuwachs größer ist als an einer anderen, dann kann man ja schon mal hinschauen: Woran liegt das? Sind die Professoren unterschiedlich gut? Oder liegt es an den Studienbedingungen?

In Abschnitt 2 werden wir zunächst zeigen, warum wir uns in erster Linie für den Wissenszuwachs und nicht für den Wissensstand interessieren. Würden alle Studienanfänger zufällig auf die Hochschulen verteilt werden und würden wir dann feststellen, dass die Studierenden nach fünf oder sechs Semestern unterschiedlich viel wissen, dann kämen nicht viele andere Gründe als die unterschiedliche Qualität der Hochschulen in Frage. Tatsächlich aber sind schon die Wissensstände der Studienanfänger recht unterschiedlich.

Abschnitt 3 ist der Kern unserer Analyse: Wir erläutern, wie wir den Wissenszuwachs messen, und wir prüfen, ob der Wissenszuwachs etwas mit der Qualität von Lehre und Forschung an den jeweiligen Hochschulen zu tun hat. Wir ziehen dazu mehr oder weniger etablierte Maße für einige Aspekte der Qualität von Hochschulen heran, gewonnen etwa aus Befragungen von Studierenden und Professoren oder dem Auszählen von Publikationen.

In Abschnitt 4 werden wir uns detaillierter mit den Studierenden wirtschaftswissenschaftlicher Studiengänge befassen. Einerseits wird untersucht, ob sich die Ergebnisse verändern, wenn die Fragen zum Bereich Wirtschaft nicht gleichgewichtet werden, sondern eine von Experten vorgeschlagene Gewichtung gewählt wird. Andererseits wird analysiert, ob der Einfluss der Lehrqualität an Universitäten und Fachhochschulen unterschiedlich ist.

In Abschnitt 5 werfen wir noch einen kritischen Blick auf die verwendeten Daten und Methoden, bevor wir in Abschnitt 6 ein kurzes Fazit ziehen.

2. Wissensunterschiede von Studienanfängern

Universitäten mit gutem Ruf gab es schon immer. Schüler waren in dieser Hinsicht aber in früheren Zeiten eher schlecht informiert. Dies gilt insbesondere für Zeiten vor der Etablierung des Internets. Heute ist das anders (vgl. auch den Beitrag von Helbig und Ulbricht in diesem Band). Rankings sind leicht verfügbar und werden in Internetforen von prospektiven Studienanfängern diskutiert. Hier drei charakteristische Beispiele:

„Laut Berichten, Rankings usw. sind ja Mannheim und München eigentlich die besten Unis für VWL. Nur hab ich gehört, dass das Unigebäude in Mannheim ziemlich runtergekommen ist und, dass in München der Kontakt zwischen den Schülern aus irgendeinem Grund nicht so gut ist. Vorurteile?“ (Beitrag vom 16 Okt 2005 auf <http://www.uni-protokolle.de/foren/viewt/38286,0.html>, Zugriff am 6.1.2010)

„Ich selbst werde wohl in Mannheim studieren, da die Uni laut Rankings (die ich nicht überbewerten will, die aber bestimmt auch einen gewissen Aussagewert haben) ganz gut sein soll, und weil das benachbarte Heidelberg ne Uni hat, die in Sachen Entwicklungshilfe ganz gut aufgestellt sein soll... (...) Ich hab auch die Möglichkeit in Nürnberg Intern.VWL zu studieren (mit 2 Fremdsprachen), bin mir aber noch nicht so ganz im Klaren über die Qualität des Studiums dort...“ (Beitrag vom 16 Okt 2005 auf <http://www.uni-protokolle.de/foren/viewt/38286,0.html>, Zugriff am 6.1.2010)

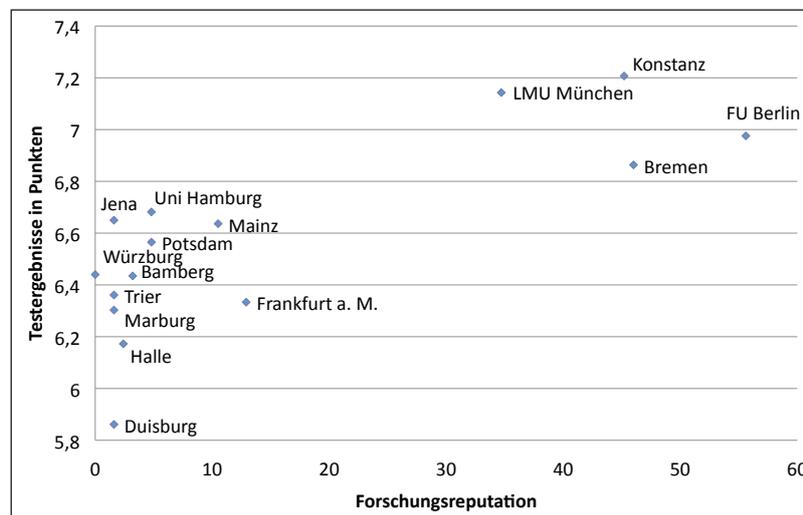
„Auch VWLer haben ihre Rankings (...) Die HU Berlin, welche dort auf #2 ist, sollte man denk ich mal auch noch erwähnen. Ich kenn sie nicht wirklich, hab aber öfters gelesen, dass sie besonders in Makro sehr gut sein soll.“ (Beitrag vom 04 Apr 2006 auf <http://www.uni-protokolle.de/foren/viewt/62635,0.html>, Zugriff am 6.1.2010)

Es gibt verschiedene Gründe, weshalb es nun gerade die guten Studienanfänger sein könnten, die sich an den „top-gerankten“ Universitäten einschreiben. Eine Möglichkeit ist, dass knappe Studienplätze insbesondere an begehrten Universitäten zu anspruchsvollen Zulassungskriterien (wie Abiturnotendurchschnitt oder auch Testergebnisse) führen. Es kann aber auch sein, dass es vor allem die besseren Schüler sind, die ein Interesse daran haben, an einer besonderen Universität zu studieren.

Glaubt man dem CHE-Ranking (Quelle: www.zeit.de/ranking), dann gibt es in Deutschland beispielsweise im Fach Politik sieben Universitäten mit hoher Forschungsreputation. Davon haben allerdings bei dreien (HU Berlin, Heidelberg und Mannheim) nicht mehr als 20 Po-

litik-Studienanfänger (Erst- und Zweitsemester) den Studententest mitgemacht, so dass sie für Abbildung 1 nicht berücksichtigt wurden. Die anderen vier, im rechten Teil von Abbildung 1 deutlich auszumachen, sind die FU Berlin, die LMU München sowie die Universitäten Bremen und Konstanz. Im Durchschnitt liegt die Leistung gut einen halben Punkt über dem Durchschnitt der Studienanfänger anderer Universitäten. Das entspricht etwa 10 Prozent. Diese Unterschiede sind nicht dramatisch, aber doch so deutlich, dass ein Teil der möglicherweise höheren Leistung der Studierenden in den höheren Semestern durch diese Selbstselektion erklärbar ist.

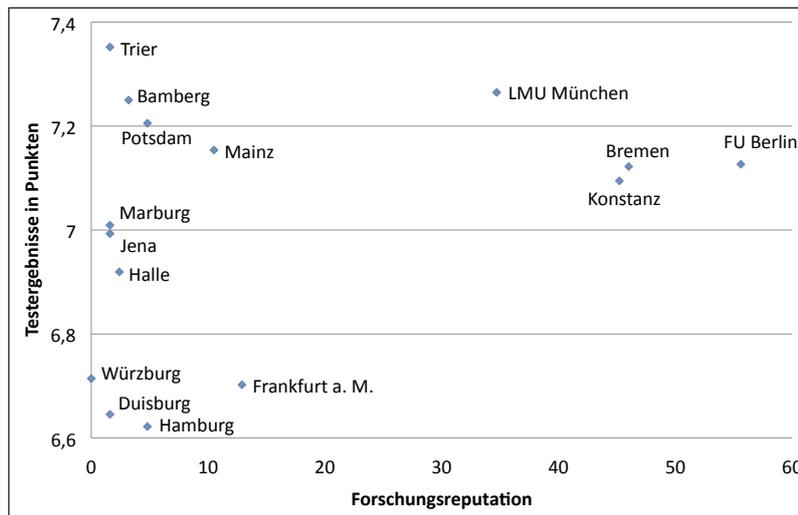
Abbildung 1: Je höher die Forschungsreputation der Universität, desto besser tendenziell das Studententest-Ergebnis der Studienanfänger (1. und 2. Semester) im Fach Politik



Einen ähnlichen Zusammenhang gibt es für die Wirtschaftswissenschaften, nicht aber für das Fach Geschichte, in dem die vom CHE ausgewiesene „Forschungsreputation“ die Universitäten wesentlich weniger differenziert. Für die kultur- und naturwissenschaftlichen sowie sonstigen Fächer können derartige Aussagen leider nicht getroffen werden, weil im Studententest keine hinreichend differenzierten Wissensgebiete formiert wurden, die mit Fächern des CHE-Kanons übereinstimmen.

Abbildung 2 zeigt nun die Leistung der höheren Semester, die zumindest teilweise unter dem Einfluss des Studiums zustande kommt. Der Vergleich zu Abbildung 1 lässt ein Aufholen der Politik-Studierenden an einigen Universitäten erkennen, genauer gesagt einen großen Abstand der Punktzahlen beim Wissenstest Politik von Fortgeschrittenen und Anfängern. Gibt es eine Erklärung dafür, dass diese Abstände an verschiedenen Universitäten unterschiedlich groß sind? Das soll im folgenden Abschnitt geklärt werden.

Abbildung 2: Zusammenhang von Forschungsreputation der Universität und Studentenpisa-Ergebnis der Fortgeschrittenen (3. Semester und höher) im Fach Politik



3. Wo wird am meisten gelernt? Und warum?

Mehr als 600.000 Personen haben am Studentenpisa-Test teilgenommen. Jede dieser Personen jedoch nur zu einem Zeitpunkt: Elftklässler in der 11. Klasse, Erstsemester im ersten Semester etc. Mit anderen Worten: Das, was uns in diesem Abschnitt interessiert, nämlich das, was eine Person hinzulernt, ist statistisch nicht erfasst. Wir greifen daher zu einer Hilfskonstruktion: Wir vergleichen für jede Hochschule die durchschnittliche Leistung der Studienanfänger mit der durchschnittlichen Leistung der Fortgeschrittenen. Wo viel gelernt wird, da sollte der Leistungsabstand zwischen Fortgeschrittenen und Anfängern besonders groß sein.

Für dieses auf Durchschnitt basierende Vorgehen zahlen wir allerdings auch einen Preis: Durch die gewählte Methode basieren unsere Analyse nicht mehr auf den originären Individualdaten, sondern auf Durchschnittswerten für die Universitäten. Dass wir dies „für jede Hochschule“ tun, müssen wir gleich relativieren. Es kommt gar nicht selten vor, dass nur ein einziger Studienanfänger einer bestimmten Fachrichtung an einer Hochschule teilgenommen hat; im Fach Politik zum Beispiel an den Universitäten Oldenburg und Siegen. Die Auswertung sollte nun nicht von den Zufällen abhängen, die die Leistung dieses einen Studierenden bestimmen, der vielleicht alkoholisiert, übermüdet oder ohne jeden Ehrgeiz teilgenommen hat; der eine oder andere wird sogar seinen Spaß daran gehabt haben, bewusst falsch zu antworten. Wir beschränken uns also auf solche Universitäten, an denen die Gruppe der Studienanfänger und der Fortgeschrittenen mit jeweils mehr als 20 Personen besetzt ist, so dass wir davon ausgehen können, dass das Testergebnis einigermaßen repräsentativ ist.

Von den fünf Wissensgebieten des Studentenpisa-Tests widmen wir Geschichte, Wirtschaft und Politik dabei besondere Aufmerksamkeit, weil hier eine klare Zuordnung von Studien-

fach und Wissensgebiet möglich ist und – wie bereits erwähnt – weitere Daten wie die CHE-Daten zur Verfügung stehen. Tabelle 1 zeigt, um wieviel das Testergebnis der Fortgeschrittenen in diesen Fächern besser ausfällt als das Testergebnis der Anfänger.

Tabelle 1: Das Abschneiden von Studierenden der Fächer Geschichte, Politik und Wirtschaftswissenschaften in ihren eigenen Wissensgebieten

Fach	Testergebnis Studienanfänger	Testergebnis Fortgeschrittene	Differenz
Geschichte	6,25	6,96	0,71
Politik	6,59	7,04	0,45
Wirtschaftswissenschaften	5,41	6,10	0,68

Anmerkungen: Die Durchschnitte basieren auf den Angaben für die Universitäten, an denen im jeweiligen Fach jede Gruppe mit mehr als 20 Personen besetzt ist. Anfänger: Erst- und Zweitsemester; Fortgeschrittene: Drittsemester und höhere Semester.

Es wäre falsch, den in Tabelle 1 dokumentierten Wissenszuwachs allein der Lehre an den Universitäten zuzuschreiben. Hierfür gibt es mehrere Gründe; der einfachste und vermutlich gewichtigste ist, dass im Studentenpisa-Test Allgemeinwissen eine große Rolle spielte. Auch wer Informatik studiert und nur hin und wieder Zeitung liest, verbessert seine Politikkenntnisse. Um einen Eindruck von der Größe dieses Effektes zu bekommen, haben wir errechnet, um wie viel Studierende der Naturwissenschaften und verwandter Fächer (Agrarwissenschaft, Architektur, Bauingenieurwesen, Biologie, Chemie, Elektrotechnik, Ernährungswissenschaft, Forstwissenschaft, Geographie, Informatik, Ingenieurwissenschaften, Maschinenbau, Naturwissenschaften, Pharmazie, Physik) und der Kulturwissenschaften (Anglistik, Germanistik, Kunst, Medienwissenschaften, Sprach- und Kulturwissenschaften) außerhalb ihrer Wissensgebiete im Laufe des Studiums hinzulernen (Tabelle 2). Vergleicht man die beiden Tabellen, so stellt man fest, dass etwa 58% des Wissenszuwachses in Geschichte und 72% des Wissenszuwachses zum Thema Wirtschaft auch bei Studierenden ganz anderer Fächer auftritt.

Tabelle 2: Das Abschneiden von Studierenden der Kultur- und Naturwissenschaften in den Wissensgebieten Geschichte, Politik und Wirtschaft

Wissensgebiet	Testergebnis Studienanfänger	Testergebnis Fortgeschrittene	Differenz
Geschichte	4,94	5,36	0,41
Politik	4,62	5,13	0,51
Wirtschaftswissenschaften	4,79	5,28	0,49

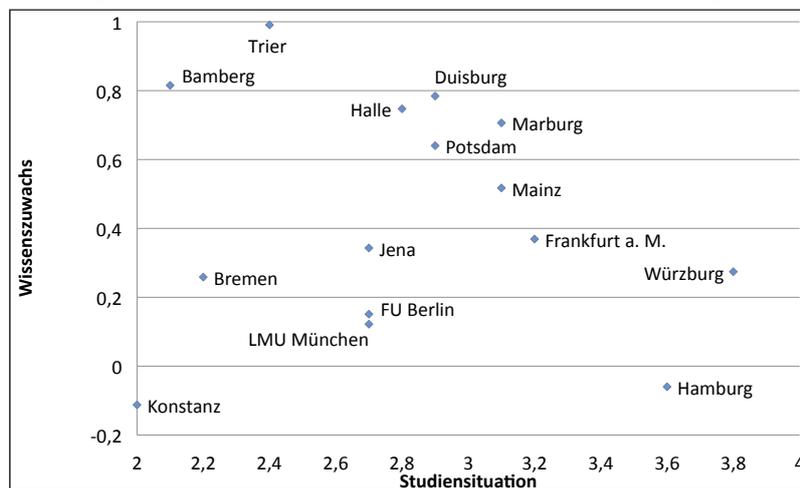
Anmerkungen: Die Durchschnitte basieren auf den Angaben für die Universitäten, deren Ergebnisse der Tabelle 1 zugrunde liegen. Anfänger: Erst- und Zweitsemester; Fortgeschrittene: Drittsemester und höhere Semester.

Damit wird die Bedeutung des Studiums für den Wissenszuwachs aber unterschätzt. Wer Geschichte zu studieren beginnt, der weiß schon mehr über Geschichte als seine Kommilitonen,

die etwas anderes studieren. Entsprechend geringer ist die Wahrscheinlichkeit, dass Informationen für ihn neu sind: Es ist daher schwerer, einen Studierenden von 7 auf 8 Punkte zu verbessern als von 4 auf 5. Dies gilt in besonderem Maße dann, wenn die Fragen unterschiedlich schwer sind und wenn das, was im Laufe des Studiums noch dazugelernt wird, gerade das ist, was sich nicht so leicht erschließt. Damit ist auch der Eindruck, dass Politikstudierende weniger hinzulernen als Kultur- und Naturwissenschaftsstudierende, trügerisch. Um dieses Problem vollständig zu lösen, müsste der Schwierigkeitsgrad der einzelnen Fragen bestimmt werden. Dies wäre z. B. durch die Schätzung von Rasch-Modellen möglich, die bei den dreijährlichen PISA-Untersuchungen der OECD zur Anwendung kommen. Dann könnten simultan die Schwierigkeit der Fragen und die Kompetenzen der befragten Studierenden geschätzt werden. Dieses Vorgehen ist mit den Studententpisa-Daten aber nicht möglich.

Wir wollen nun prüfen, ob ein signifikanter Zusammenhang besteht zwischen dem Lernen an den Universitäten, wie wir es in Tabelle 1 mit einem zugegebenermaßen groben Maß erfassen, und anderen Variablen, die die Qualität von Forschung und Lehre an den Universitäten erfassen. Abbildung 3 zeigt ganz grob das Prinzip. Auf der Ordinate (vertikale Achse) ist die Differenz zwischen den Punktzahlen der Fortgeschrittenen und der Anfänger (Wissenszuwachs_Politik) abgetragen. Auf der Abszisse (horizontale Achse) ist die Bewertung der Studiensituation insgesamt durch die Studierende der jeweiligen Universität, skaliert als Schulnote, abgetragen (Studiensituation_Politik_CHE). Erhoben wird dieser Indikator vom Centrum für Hochschulentwicklung (CHE).

Abbildung 3: Beispiel Politik: Bessere studentische Benotungen der Studiensituation gehen tendenziell mit höherem Wissenszuwachs (Differenz des Studententpisa-Ergebnisses zwischen höheren Semestern und Studienanfängern) einher



Mit Hilfe der Regressionsanalyse versucht man, durch die in Abbildung 3 dargestellte Punktwolke eine Gerade zu legen, so dass die Summe der quadrierten (vertikalen) Abstände zwischen den einzelnen Punkten der Punktwolke und der Regressionsgeraden minimal wird. In diesem Fall ergibt sich:

$$\text{Wissenszuwachs_Politik} = a_0 - a_1 \cdot \text{Studiensituation_Politik_CHE}$$

Dabei sind a_0 und a_1 die geschätzten Regressionskoeffizienten, die wie folgt zu interpretieren sind: Verbessert sich eine Universität in Bezug auf die Studiensituation im Fach Politik z.B. von der Note 3 auf die Note 2, dann erwarten wir eine Erhöhung des Lernfortschritts zwischen den ersten beiden und den höheren Semestern um a_1 .

Diese Analyse greift aber noch zu kurz. Wir müssen gleichzeitig berücksichtigen, dass unser Maß für den Wissenszuwachs Wissenszuwachs_Politik auch vom Ausgangsniveau abhängt. Dieses nennen wir Anfänger_Score_Politik. Außerdem untersuchen wir Folgendes: Ist der Lernfortschritt vielleicht höher dort, wo besser geforscht wird? Auch hierzu wird ein Indikator vom CHE erhoben: Professoren an deutschen Universitäten werden gebeten, fünf Universitäten in Deutschland zu nennen, die sie für „führend“ in der Forschung in ihrem Fach halten. Für jede Universität wird ausgewiesen, wie viel Prozent der Professoren sie für „führend“ halten. Für jede Universität wird ausgewiesen, wie hoch der Anteil der Professoren war, der sie genannt hat.

Die geschätzte Regressionsgleichung, die viel zuverlässiger ist als die notwendigerweise auf zwei Variablen beschränkte Abbildung 3, sieht dann so aus:

$$\begin{aligned} \text{Wissenszuwachs_Politik} = & 5,6386 \\ & - 0,3300 \cdot \text{Studiensituation_Politik_CHE} \\ & - 0,6376 \cdot \text{Anfänger_Score_Politik} \\ & - 0,0054 \cdot \text{Forschungsreputation_Politik_CHE} \end{aligned}$$

Das Ergebnis (für Details siehe Tabelle 3 im Anhang) ist wie folgt zu interpretieren: Die studentische Benotung der Studiensituation hat den erwarteten Einfluss auf den Wissenszuwachs: Verbessert sich die Note z. B. von 3 auf 2 oder von 2 auf 1, erhöht sich der Wissenszuwachs um 0,33. Steigt die Forschungsreputation, so *sinkt* der Wissenszuwachs – ein unerwartetes Ergebnis, das allerdings nicht statistisch signifikant ist, so dass wir diesen negativen Einfluss nicht als gesichert annehmen dürfen. In jedem Fall deuten unsere Ergebnisse darauf hin, dass die Forschungsstärke einer Universität unerheblich für den mit Studententpisa gemessenen Wissenszuwachs ist. Dies gilt auch für die anderen Fächer, die wir näher betrachten, d.h. für Wirtschaft und Geschichte.

Man hätte auch ein anderes Ergebnis erwarten können, und zwar aus folgenden Gründen:

1. Universitäten mit Professoren mit hoher Forschungsreputation könnten eine besondere Faszination auf die Studierende ausüben, was ein didaktischer Vorteil sein könnte.
2. Sie könnten in der Lage sein, den Stoff besser zu vermitteln, da sie in der Vorlesung gar nicht mehr über die Inhalte nachdenken müssen, sondern nur noch über die Präsentation.
3. Ihre Forschungsreputation zieht besonders fähige Mitarbeiter an, die wiederum bessere Übungen etc. anbieten können als der Durchschnittsassistent.

All dies mag in einigen Fällen gelten, aber insgesamt finden wir dafür keine Anhaltspunkte. Möglicherweise dominieren andere Effekte:

1. Universitäten oder Fachbereiche, die wissen, dass sie keine Spitzenforschung hervorbringen, haben größere Anreize, sich in der Lehre zu profilieren.
2. Professoren, die in der Spitzenforschung ohnehin nicht recht mitmischen können, geben vielleicht leichter einen Teil ihrer Zeit für die Vorlesungsvorbereitung auf.

Die vorliegenden Daten erlauben allerdings leider keine differenzierte empirische Überprüfung dieser Thesen. Folglich kann aus unserer Sicht nur der Schluss gezogen werden: Wo besser geforscht wird, wird nicht gleich auch besser gelehrt.

Und wo wird nun besser gelehrt? Während wir im Fach Politik bei Universitäten, die nach den CHE-Daten bessere Studienbedingungen aufweisen, einen höheren Wissenszuwachs fanden, gilt dies nicht für Geschichte: Tabelle 3 im Anhang zeigt, dass wir hier keinen signifikanten Einfluss der studentischen Benotung der Studienbedingungen finden. Beim Fach Wirtschaft ist das Ergebnis gemischt (Tabelle 4 im Anhang). Hier hängt das Ergebnis davon ab, ob wir das CHE-Ranking einsetzen oder das aktuellste Focus-Ranking aus dem Jahr 2007, das ordinalskaliert mit drei Ausprägungen (Spitzengruppe, Mittelgruppe, Schlussgruppe) jeweils für Forschung und Lehre für mehrere Fächer vorliegt. Die Korrelation zwischen der Lehrbewertung durch das CHE und der durch Focus ist ziemlich gering, so dass im Prinzip andere Ergebnisse nicht auszuschließen sind. Und genau das ist passiert (Tabelle 4 im Anhang): Die Bewertung der Studiensituation einer Universität im CHE-Ranking korreliert nicht mit dem Wissenszuwachs im Studentenpisa-Test, während das Focus Ranking einen (schwach) signifikanten Einfluss aufweist.

Allerdings sind die Ergebnisse mit Vorsicht zu interpretieren. Im nächsten Abschnitt zeigen wir, wo Probleme in Daten und Methoden liegen.

4. Ergänzende Analysen zu den Wirtschaftswissenschaften

4.1 Die Bewertung des Wissens: Ergänzungen zum Wissensgebiet Wirtschaft

Warum gibt es Studentenermäßigungen für Kinokarten, nicht aber für Brötchen oder Fahrräder? Warum werden manche Produkte zwar weltweit angeboten, aber lokal konzentriert produziert, wie Messer in Solingen oder Filme in Hollywood? Warum zahlen Blumenläden in Einkaufszentren eine geringere Miete als Schuhläden, obwohl sie in Einkaufsstraßen dieselbe Miete zahlen? Vorrangiges Ziel von Ökonomievorlesungen ist nicht, dass Studierende die konkreten Antworten auf diese Fragen lernen. Es geht vielmehr darum, sie mit einem methodischen Instrumentenkasten auszustatten, der es ihnen erlaubt, für diese und ähnlich gelagerte Fragen Antworten selbst zu finden. Ob das gelingt, ist mit einem online durchgeführten Multiple-Choice-Test nur bedingt festzustellen.

Die Studentenpisa-Fragen zum Thema Wirtschaft konnten sich nicht auf die Methoden beziehen, die in wirtschaftswissenschaftlichen Lehrveranstaltungen vermittelt werden. Vielmehr stellen die Fragen auf Kenntnisse über den „Gegenstandsbereich Wirtschaft“ ab. Aber die Wichtigkeit der abgeprüften Kenntnisse ist von Frage zu Frage sehr unterschiedlich. Wer weiß, ob Time Warner von AOL, Google, MySpace oder Lycos übernommen worden war, ist deshalb noch kein besserer Ökonom als jemand, der das nicht weiß. Und dieses Wissen würde

einem Prüfling wohl in keiner Ökonomie-Klausur an einer deutschen Universität etwas nützen. Auch nach der offiziellen Zahl der Arbeitslosen wird zwar im Studententpisa-Test, aber nur selten direkt in Prüfungen an Universitäten gefragt, wenngleich sie wohl in jeder Vorlesung zur Makroökonomik vorkommt. Wir wollten uns, was die Wichtigkeit der Fragen betrifft, aber nicht auf unser eigenes Bauchgefühl verlassen, sondern haben sechs Experten (zwei Professoren, zwei promovierte und zwei nicht promovierte Ökonomen), die in sehr unterschiedlichen wirtschaftswissenschaftlichen Feldern tätig sind, für jede der auf die Wirtschaft bezogenen Studententpisa-Fragen um zwei Angaben gebeten:

1. „Wichtigkeit“: Wie wichtig ist es für Studierende der Wirtschaftswissenschaften, die Antwort auf die Frage zu kennen? Hierzu war mit einer Zahl von 1 (entspricht völlig unwichtig“) bis 10 (entspricht „sehr wichtig“) zu antworten. (Kendalls Konkordanzkoeffizient, der die Übereinstimmung der Beurteiler auf einer Skala von 0 bis 1 misst, nimmt hier den Wert 0,67 an.)
2. „Hochschulrelevanz“: Sollte dieses Wissen im Studium der Wirtschaftswissenschaften vermittelt werden? Hier waren drei Antworten möglich: „Ja, unbedingt“, wofür wir die Zahl 1 vergeben haben, „vielleicht/weiß nicht“, wofür wir die Zahl 0,5 vergeben haben, und „nein, ist nicht unsere Aufgabe“, wofür die 0 vergeben wurde (Kendalls Konkordanzkoeffizient = 0,69).

Aus der Multiplikation der Wichtigkeit und der Universitätsrelevanz errechneten wir dann Gewichte für die Fragen, und zwar so, dass Teilnehmer, die alle neun Fragen zum Thema Wirtschaft richtig beantworteten, genau 9 Punkte bekamen.

Dass Wichtigkeit und Universitätsrelevanz nicht dasselbe ist, zeigt sich etwa bei der Frage nach den Google-Gründern. Im Durchschnitt vergaben unsere sechs Juroren hier eine 5 für die Wichtigkeit, im Vergleich zu 6,1 für den Durchschnitt aller Wirtschaftsfragen. Hier ist die Frage also durchaus noch im Rahmen, bei der Universitätsrelevanz fiel sie glatt durch, sie wurde von allen Juroren verneint, so dass die Frage insgesamt das Gewicht 0 erhält.

Einige andere Fragen sind nach Meinung unserer Juroren sowohl unwichtig (Wichtigkeit < 3) als auch in keinster Weise universitätsrelevant: René Obermann, Klaus Zumwinkel oder das Commerzbank-Logo auf einem Foto erkennen oder den Namen der ersten Internet-Suchmaschine kennen. Die höchsten Gewichte insgesamt erhielten, weil sie sowohl wichtig als auch universitätsrelevant sind, die Fragen nach der Definition des Bruttoinlandsproduktes (BIP) und dem deutschen BIP pro Kopf; ein mindestens doppelt so hohes Gewicht wie eine durchschnittliche Frage erhielten auch die Fragen nach dem Zweig der gesetzlichen Sozialversicherung mit dem höchsten Beitragssatz, nach dem Ökonomen, auf den sich Verfechter staatlicher Konjunkturprogramme berufen, nach dem Hartz IV-Regelsatz, nach der Tätigkeit von Rating-Agenturen und nach dem Verhältnis von Umsatz und Gewinn.

Exemplarisch für den Bereich Wirtschaft konnten wir zusätzlich zu der Zahl richtiger Antworten nun also eine gewichtete Zahl richtiger Antworten für jeden Teilnehmer berechnen. Diese beiden Variablen korrelieren sehr stark miteinander ($r > 0,8$). Dazu passt, dass alle Ergebnisse, über die wir im vorherigen Abschnitt berichtet hatten, bestehen bleiben, wenn wir die ungewichtete durch die gewichtete Punktzahl bei den Wirtschaftsfragen ersetzen.

4.2 *Lernen an Fachhochschulen und Universitäten: Gibt es einen Unterschied?*

Was unterscheidet Fachhochschulen von Universitäten? Unter anderem wohl dies: In Fachhochschulen spielt die Vermittlung von konkretem, anwendbarem Wissen eine größere Rolle, weniger wichtig als an Universitäten ist dagegen die Vermittlung der Forschungsmethoden, mit denen neues Wissen überhaupt erst generiert wird. Wenn das stimmt, dann hat der Studentenpisa-Test mehr mit dem zu tun, was FH-Studierende lernen sollten, als mit dem, was für ihre Kommilitoninnen und Kommilitonen an den Universitäten im Vordergrund steht. Im Vergleich zu Universitäten sollte die Qualität der Lehre an Fachhochschulen daher einen größeren Einfluss auf die Leistung im Studentenpisa-Test haben. Wir finden in der Tat, dass dies der Fall ist. Verbessert eine Universität ihre Lehrbewertung im Focus-Ranking von 1 auf 3, also von der schlechtesten auf die beste Bewertung, so ist nach Tabelle 4 im Anhang eine Erhöhung der Studentenpisa-Punktzahl um 0,17 zu erwarten. Das ist wenig in Relation zu FHs, wo schon die Verbesserung um eine einzige Schulnote in der CHE-Bewertung der „Studiensituation insgesamt“ einen höheren Einfluss hat (an Universitäten ist gar kein statistisch signifikanter Einfluss dieser Variable festzustellen). Für Politik und Geschichte können keine vergleichbaren Untersuchungen durchgeführt werden, da keine entsprechenden Studiengänge an Fachhochschulen angeboten werden.

5. Ein kritischer Blick auf die verwendeten Daten und Methoden

Die Generierung und Aussagekraft der Studentenpisa-Testergebnisse wird an anderer Stelle in diesem Band diskutiert. Wir beschränken uns daher auf kurze Bemerkungen zu den von uns übernommenen Daten über Lehre und Forschung (5.1) und zu unserer Methode (5.2).

5.1 *Die Bewertung von Lehre und Forschung*

Das CHE stellt für ca. 20 Fächer eine ganze Reihe von Einzelindikatoren zur Verfügung. Bisher haben wir zwei davon verwendet: Die Forschungsreputation und die Studiensituation insgesamt. Die Arbeit mit studentischen Lehrevaluationen ist in ökonomischen Untersuchungen durchaus üblich (Soo, 2009). Die Messung und Evaluierung von Forschungsleistung durch das CHE wurde sehr kritisch diskutiert (z. B. Ursprung, 2003). Für unsere Zwecke ist diese – an sich sehr überzeugende – Kritik aber nicht unmittelbar relevant, da es uns in insbesondere in Abschnitt 2 gerade nicht um die „wahre“ Forschungsqualität ging, sondern um die öffentlich wahrgenommene Forschungsqualität, die durch das CHE mitgeprägt wird. Aber warum nicht noch mehr CHE-Indikatoren berücksichtigen? Einige von ihnen könnten durchaus einen Einfluss auf das Lernen an der Universität haben, zum Beispiel die „Reputation in Studium und Lehre“, die ähnlich wie die Forschungsreputation bei Professoren erfragt wurde. Allerdings ist auch das Ergebnis extrem ähnlich: Die „Reputation in Studium und Lehre“ und die „Forschungsreputation“ sind so stark miteinander korreliert (z. B. BWL: $r=0,95$; VWL: $r=0,98$), dass man die eine gegen die andere Variable austauschen könnte, ohne dass dies zu spürbar anderen Ergebnissen führt.

Fast ebenso groß ist dieses Problem für die studentischen Benotungen der Studienbedingungen: auch hier gibt es Alternativen wie Benotungen der Betreuung oder des Lehrangebots, aber diese Benotungen sind ebenfalls hoch korreliert und somit (ohne entscheidende Ände-

rung des Ergebnisses) gegeneinander austauschbar, aber nicht gemeinsam in einer Regressionsgleichung einsetzbar.

5.2 Die Methode

Rufen wir uns noch einmal in Erinnerung, dass wir mit Fortgeschrittenen und Anfängern verschiedene Gruppen vergleichen. Die Fortgeschrittenen waren einmal Anfänger und haben hinzugelern. Allerdings könnte sich die Zusammensetzung der Gruppen im Laufe der Zeit ändern. Wenn insbesondere wissenschwache Studierende ihr Studium nach einem oder zwei Semestern beenden, dann führt dies zu einem besseren Abschneiden der Fortgeschrittenen im Vergleich zu den Studienanfängern. Zudem ist plausibel anzunehmen, dass die schwachen Studierende gerade dort aufgeben, wo die Bedingungen für die Betreuung schlecht sind. Anders formuliert: Wo die Betreuung gut ist, werden schwache Studierende an der Universität gehalten. Bei der Methode, die wir verwenden, führt das dazu, dass wir den besonders interessanten Effekt *unterschätzen*, nämlich den Einfluss der Lehrqualität auf den Wissenszuwachs.

Abgesehen davon, dass Studierende das Studium abbrechen, können auch Studienortwechsel der Studierenden eine Rolle spielen. Bietet eine Universität ein besonders attraktives Masterprogramm, dann kann sie sich die besten Bachelor-Absolventen aussuchen. Dass die hohen Semester dann viele Punkte erzielen, hängt aber nicht mit der Qualität der Lehre in den ersten Semestern zusammen. Daher haben wir die bisher präsentierten Regressionsansätze variiert. Als Gruppe der Fortgeschrittenen haben wir alternativ nur die Studierenden des 5. und 6. Semesters gewählt. Obwohl dann die angesprochenen Wanderungseffekte ausfallen, zeigt sich nicht das erwartete Ergebnis, d. h. der Einfluss der Lehrqualität ist nicht deutlicher oder größer als mit der in Abschnitt 3 und 4 verwendeten Methode. Zwei Gründe spielen eine Rolle: Einerseits sind die Studierenden im 5. und 6. Semester weniger stark vertreten als die Anfänger (1. und 2. Semester), und ex definitione ist diese Gruppe viel kleiner als die ursprüngliche Gruppe der Fortgeschrittenen (3. und höheres Semester), während die Streuung der Durchschnittswerte für die einzelnen Universitäten der jeweiligen Fächer z. T. deutlich höher ist. Zufallseinflüsse spielen somit eine größere Rolle. Ein weiterer Grund könnte daran bestehen, dass Wanderungseffekte nur bei einigen Universitäten zu beobachten sind. Dies könnte Gegenstand weiterer Forschungsarbeiten sein, die das Zuspielen weiterer Daten notwendig machen würde.

Ein Nachteil unserer Methode, auf den wir schon hingewiesen hatten, ist, dass wir auf die Informationen über die einzelnen Studierenden verzichten müssen, damit wir die Wissenszuwächse untersuchen können. Eine Möglichkeit, wie wir unsere Frage aber auch mit den Mikro-Daten über die Punktzahlen der einzelnen Teilnehmer angehen könnten, skizzieren wir für spezialisierte Leser in Anhang 2 – es bedarf, kurz gesagt, eines leider erst zur Hälfte ausgearbeiteten Tricks, der zudem das oben erläuterte Problem, dass sich die Zusammensetzung der Studentenschaft von Semester zu Semester ändert, auch nicht löst. Schlussfolgerungen lässt zur Zeit nur unser Querschnittsvergleich des Lernens an verschiedenen Hochschulen zu.

6. Schlussfolgerungen

Die Studententypen-Daten bieten eine hervorragende Möglichkeit, den Wissenstand von Studierenden zu untersuchen; über Umwege aber auch die Wissenszuwächse. In dem vorliegen-

den Beitrag wurden erste Ergebnisse präsentiert. Weitere Analysen, insbesondere auch in Zusammenspiel mit weiteren Daten zur Qualität in Lehre und Forschung, könnten z. B. auch darlegen, welche Wissenszuwächse durch ein Studium in fachfremden Wissensgebieten generiert werden können oder eben gerade nicht, aufgrund des vielbeklagten „Spezialistentums“.

Im Hinblick auf die eingangs formulierten Fragestellungen können wir uns – bei aller Vorsicht – eines Ergebnisses sicher sein: Das „~~gegen~~^{oder}“ im Titel reflektiert zwar unsere Frage, wird den Ergebnissen aber nicht ganz gerecht. Denn es gilt beides: Auf das Studententypa-Ergebnis hat zunächst einmal die Verteilung der Schüler auf die Hochschulen einen spürbaren Einfluss. Deutlich erkennbar ist, dass forschungsstarke Universitäten gute Studierende anziehen und/oder auswählen (Abschnitt 2). Wie stark die Studierenden sich danach verbessern, hängt davon ab, wie gut die Lehre ist. Dieses Ergebnis erzielen wir in Abschnitt 3 dann, wenn wir ein Maß für die Lehrqualität verwenden, das das Urteil der Studierenden abbildet. Professoren dagegen scheinen zu glauben, dass die Lehre genau dort gut ist, wo die Forschung gut ist. Wir fanden allerdings, dass die Qualität der Forschung nicht auf die Wissensvermittlung durchschlägt.

Literatur

- Ai, C. & Norton, E.C. (2003). Interaction terms in logit and probit models, *Economics Letters*, 80, 123–129.
- Soo, T. K. (2009). Estimating the production function of university students, mimeo: <http://www.lums.lancs.ac.uk/publications/viewpdf/006023/>
- Ursprung, H. (2003). Schneewittchen im Land der Klapperschlangen: Evaluation eines Evaluators. *Perspektiven der Wirtschaftspolitik*, 2, 177-189.

Anhang 1: Regressionsergebnisse*Table 3:* Einflüsse auf den Wissenszuwachs^a in Politik (Spalten 1 und 2) und Geschichte (Spalte 3)

	Politik		Geschichte
	(1)	(2)	(3)
Punktzahl Studienanfänger ^b	-0,638** (3,13)	-0,928** (2,36)	-0,790*** (4,39)
Forschungsreputation ^c	-0,005 (1,33)	0,002 (0,33)	0,003 (0,42)
Studiensituation nach CHE ^c	-0,330*** (3,22)		0,019 (0,08)
Studiensituation nach Focus ^c		-0,072 (0,53)	
Konstante	5,64*** (4,28)	6,71** (2,54)	5,95*** (5,56)
n	15	13	16
R ²	0,79	0,65	0,63

Anmerkungen:

a: Durchschnittliche Punktzahl der Studierenden an einer Universität des dritten und höherer Semester abzüglich der durchschnittlichen Punktzahl der Erst- und Zweitsemester der Universität

b: Studierende der Politikwissenschaften (Spalten 1 und 2) bzw. Geschichte (Spalte 3)

c: In Politik (Spalten 1 und 2) bzw. Geschichte (Spalte 3)

(t-Werte in Klammern)

Table 4: Einflüsse auf den „Wissenszuwachs“ in Wirtschaft

	Universitäten		Fachhochschulen
	(4)	(5)	(6)
Punktzahl Studienanfänger ^b	-0,306*** (3,46)	-0,410*** (4,81)	-0,681*** (9,47)
Forschungsreputation BWL	0,003 (1,34)	0,002 (1,11)	
Studiensituation BWL nach CHE	0,019 (0,24)		0,221*** (2,86)
Studiensituation BWL nach Focus		0,083* (1,78)	
Konstante	2,29*** (4,03)	2,75*** (6,28)	4,56*** (10,46)
n	55	56	98
R ²	0,22	0,31	0,49

Anmerkungen:

a: Durchschnittliche Punktzahl der Studierenden an einer Hochschule des dritten und höherer Semester abzüglich der durchschnittlichen Punktzahl der Erst- und Zweitsemester der Hochschule

b: Studierende der Fächer BWL, VWL, Wirtschaftsingenieurwesen etc.

(t-Werte in Klammern)

Anhang 2: Zur Verwendbarkeit der Mikrodaten

Eine Möglichkeit, wie wir unsere Frage auch mit den Mikro-Daten über die Punktzahlen der einzelnen Teilnehmer angehen könnten, obwohl jeder nur zu genau einem Zeitpunkt am Studentenpisa-Test teilnimmt, könnte so funktionieren: Wir schätzen zunächst den Einfluss der Semesterzahl auf das Wissen. Im Durchschnitt gilt, dass das Wissen zunimmt, je höher die Semesterzahl ist, und dies gilt – was angesichts der Studentenpisa-Fragen vielleicht etwas erstaunlich ist – noch weit über das 5. Semester hinaus. Wie dem auch sei, wir können nun auch feststellen, ob sich die Stärke dieses Zusammenhangs zwischen den Universitäten unterscheidet. Diese Methode beruht auf der – plausiblen – Annahme, dass dort, wo die Semesterzahl einen hohen Einfluss auf das Wissen hat, gut studiert wird. Die Verwendbarkeit von Interaktionsvariablen (d.h.: hängt der Einfluss der einen Variable von der Ausprägung der anderen Variable ab?) steht aber sehr in Frage: Die abhängige Variable, die Zahl der richtig beantworteten Fragen, hat nur 10 mögliche Ausprägungen, so dass ein *ordered probit*-Modell geschätzt werden müsste. Interaktionsvariablen aber verschließen sich schon bei einfachen Probit-Modellen der üblichen Interpretation (Ai & Norton, 2003).