

Klaus Wohlrabe, Sabine Gralka und Lutz Bornmann

Zur Effizienz deutscher Universitäten und deren Entwicklung zwischen 2004 und 2015

Internationale Universitätsrankings fokussieren in der Regel auf den Output von Institutionen, wie z.B. die Anzahl und Wirkung der Publikationen sowie die Anzahl renommierter Preise. Da der Output maßgeblich vom Input abhängt, vor allem von den Ausgaben und vom Personal, erscheint es sinnvoll, die Effizienz von Universitäten in den Blick zu nehmen. Die vorliegende Studie zeigt, wie sich die Effizienz deutscher Universitäten zwischen 2004 und 2015 entwickelt hat. Für die empirische Analyse wird als gängiger methodischer Ansatz die Data Envelopment Analysis (DEA) herangezogen. Mittels mathematischer Optimierung wird aus verschiedenen In- und Outputs ein relatives institutionelles Effizienzmaß bestimmt. Als Variablen für den Input werden die Ausgaben und das Personal einer Universität verwendet; als Variablen für den Output gehen die Anzahl der hochzitierten Publikationen sowie die Anzahl der Graduierten in die Analyse ein. Neben allgemeinen Trends in der Effizienz deutscher Universitäten wird in der Studie auch die individuelle Entwicklung einzelner Hochschulen beleuchtet. Generell deuten die Ergebnisse auf ein ähnliches und zeitlich konstantes Effizienzniveau in der deutschen Universitätslandschaft hin.

Seit der Publikation der ersten Universitätsrankings vor etwa 20 bis 30 Jahren spielen nationale und internationale Universitätsrankings eine immer größere Rolle in der öffentlichen Diskussion über die Leistungsfähigkeit des Wissenschaftssektors in Deutschland. Über die Ergebnisse der Rankings wird versucht zu zeigen, auf welchem Level sich die Universitäten in Deutschland bewegen – national und international gesehen. Die Veröffentlichung von neuen Ergebnissen durch wichtige Anbieter dieser Rankings (z.B. Times Higher Education oder Shanghai Ranking Consultancy) führt in der Regel zu einer Vielzahl von Pressemitteilungen, bei denen erfolgreiche Universitäten darauf hinweisen, um wie viele Rangplätze sie sich im Vergleich zum vorherigen Ranking verbessert haben. Allerdings sind deutsche Universitäten in den internationalen Rankings selten auf den ersten Plätzen zu finden. Ein wichtiger Grund für die Etablierung der Exzellenzinitiative in Deutschland war deshalb der Wunsch der Bundesregierung, es einigen Universitäten (finanziell) zu ermöglichen, zu internationalen Leuchttürmen der Forschung zu werden. Heutzutage werden die Ergebnisse der Universitätsrankings in ganz unterschiedlichen Bereichen eingesetzt: So verwenden Student*innen sie beispielsweise als Entscheidungshilfe für den Studienort, und Nachwuchswissenschaftler*innen wählen auf ihrer

Grundlage ein geeignetes Institut für die weitere akademische Laufbahn (vgl. Hazekorn 2015).

Universitätsrankings fokussieren in der Regel auf den Output von Universitäten, wie z.B. die Anzahl der Publikationen und angesehenen Preise. Dabei wird aber außer Acht gelassen, dass der Output maßgeblich vom Input bestimmt wird. Über das sogenannte Excellence-Mapping-Tool, das die Performance von Universitäten und außeruniversitären Forschungseinrichtungen weltweit visualisiert (siehe <http://www.excellencemapping.net>), kann beispielsweise der maßgebliche Einfluss der finanziellen Mittel eines Landes auf die Performance der Einrichtungen in diesem Land aufgezeigt werden (vgl. Bornmann et al. 2014). In der vorliegenden Studie beschäftigen wir uns deshalb mit der Frage, wie gut Universitäten in Deutschland ihren Input (wie z.B. das Budget) in entsprechenden Output (wie z.B. die Anzahl an Absolventen) umwandeln. Der vorliegende Artikel wirft einen Blick auf die Entwicklung der Effizienz von 70 Universitäten über einen Zeitraum von mehr als zehn Jahren.

LITERATUR

Ein umfassender Überblick über die Literatur zur Effizienzmessung findet sich für den Bildungssektor in De Witte und López-Torres (2017) und für den Hoch-

schulbereich in Rhaïem (2017) und Gralka (2018a). Die Überblicksarbeiten zeigen, dass vor allem zwei Methoden für die Effizienzmessung im Hochschulsektor verwendet werden: die in der vorliegenden Studie verwendete Data Envelopment Analysis (DEA, ein nicht parametrischer Ansatz) und die Stochastic Frontier Analyse (SFA, ein parametrischer Ansatz). Beide Methoden werden dafür eingesetzt, die Leistung von Universitäten innerhalb eines Landes zu bewerten. Neue Datenquellen, die insbesondere auf der europäischen Ebene zur Verfügung stehen, würden zwar die Evaluierung über Ländergrenzen hinweg ermöglichen; das grundsätzliche Problem der fehlenden Vergleichbarkeit der für die Analyse notwendigen Daten (vor allem Input-Daten) bleibt jedoch trotz der einheitlichen Datenquelle bestehen.

Wie die bisherigen Studien zur Effizienz zeigen, die sich auf ein einzelnes Land beziehen, sind die in den Studien verwendeten Variablen und Daten recht heterogen. Bei der Auswahl der Daten für die Analysen spielte häufig deren Verfügbarkeit eine wichtige Rolle: In vielen Studien werden nicht diejenigen Daten verwendet, die am besten geeignet sind, sondern es wird versucht, einen Kompromiss zwischen Eignung und Verfügbarkeit zu finden. Basierend auf den in der Vergangenheit publizierten Studien zur Effizienzmessung hat Gralka (2018b) ein Kernmodell formuliert, das gut verfügbare und deshalb in den bisherigen Studien häufig verwendete Variablen umfasst. Wie das Kernmodell zeigt, werden bei der Effizienzmessung zumeist die finanzielle Ausstattung, die Anzahl der Studenten und das Personal als Input, den Lehr Tätigkeiten (repräsentiert durch die Anzahl der Graduierten) und den Forschungsaktivitäten (repräsentiert durch eingeworbene Forschungsmittel und die Anzahl an Publikationen) als Output gegenübergestellt. Indikatoren, die Hinweise auf die Qualität der Forschung geben wie z.B. Zitate, wurden in den Studien – aufgrund der fehlenden Verfügbarkeit – zumeist nicht verwendet. Ausnahmen sind die Studien von Thursby (2000), Bonaccorsi et al. (2006), Agasisti et al. (2012) und Gralka et al. (2019). Zitationsbasierte Indikatoren sind jedoch wichtig in der Forschungsevaluation, da die Qualität der Forschung eine maßgebliche Größe in der Wissenschaft ist: Wenn es in der Forschungspolitik um die Performance von wissenschaftlichen Einrichtungen geht, kann man davon ausgehen, dass die Exzellenz der Einrichtungen in irgendeiner Form thematisiert wird. In der vorliegenden Studie haben wir deshalb zitationsbasierte Indikatoren verwendet, die einen Einblick in die Qualität der universitären Forschung geben können. Auch wenn Indikatoren, welche die Qualität der Lehre messen, von gleicher Relevanz wären, bietet die aktuelle Datenlage dafür bisher noch keine zufriedenstellende Lösung an.

Die zunehmende Bedeutung der Leistungsmessung im deutschen Hochschulsektor spiegelt sich durch eine gestiegene Anzahl an Studien in diesem Bereich wider. Eine der ersten Untersuchungen zur Effizienz deutscher Hochschulen wurde von Warning (2007) basierend auf Querschnittsdaten publiziert. Kempkes und Pohl (2010) veröffentlichten einige Jahre später eine Effizienzanalyse auf der Basis von Paneldaten, die zeigte, dass sich die deutschen Universitäten generell auf einem hohen Effizienzniveau bewegen. In späteren Studien wurden die Heterogenität von Institutionen berücksichtigt (vgl. Johnes und Schwarzenberger 2011) und Skalenvorteile (vgl. Olivares und Wetzel 2014) sowie persistente Faktoren (vgl. Gralka 2018a) betrachtet. Gralka et al. (2019) beschäftigten sich mit der Frage, welcher Forschungsoutput bei der Evaluation des deutschen Hochschulsektors berücksichtigt werden sollte. Vergleichsstudien zwischen deutschen Universitäten und Hochschulen eines anderen Landes lieferten Agasisti und Pohl (2012) sowie Agasisti und Gralka (2019) – in beiden Fällen wurde der italienische Hochschulsektor als Vergleich herangezogen. Aufgrund der hohen Relevanz innerhalb des deutschen Hochschulsektors wurden in weiteren Studien die im Jahr 2005 gestartete Exzellenzinitiative ebenso wie die dabei ausgezeichneten Universitäten evaluiert (vgl. Gawellek und Sunder 2016; Wohlrabe et al. 2019). Gnewuch und Wohlrabe (2018) widmeten sich in einer kürzlich veröffentlichten Studie der Effizienz von Wirtschaftsfakultäten in Deutschland.

METHODEN

In Anlehnung an die Produktionstheorie wird bei der Effizienzanalyse ein Produktionsprozess unterstellt, der sich durch Inputs und damit erzeugte Outputs beschreiben lässt. Bei der Effizienzanalyse wird eine Gruppe von Universitäten miteinander verglichen, so dass das abschließende Effizienzmaß als relativer Wert verstanden werden kann. Es wird angenommen, dass die evaluierten Universitäten mit den zur Verfügung stehenden Mitteln versuchen, den maximalen Output zu erzielen (es wird also eine Output-Orientierung unterstellt). Für eine Universität liegt der Zustand einer Dominanz vor, wenn neben ihrer eigenen Input-Output-Kombination keine Kombination seitens einer anderen Universität existiert, die mit gleichem Input einen größeren Output erzeugt.

Die Grundidee der DEA geht auf Charnes et al. (1978) zurück. Als nicht parametrisches Verfahren lässt die Methode den Vergleich mehrerer Input- und Output-Indikatoren zu und benötigt keine Annahmen hinsichtlich der zugrunde liegenden Produktions- oder Kostenfunktionen. Für die Berechnung der DEA in dieser Studie wird die Spezifikation von Banker, Charnes und Cooper (1984) verwendet –

unter der Annahme variabler Skalenerträge und eines Output-orientierten Modells (siehe oben). In Wohlrabe et al. (2019) wird eine DEA beispielhaft illustriert. Bei der Interpretation der Ergebnisse einer DEA muss berücksichtigt werden, dass die Methode einen relativen Vergleich der in die Studie einbezogenen Universitäten erlaubt. Das Ergebnis ist also kein absoluter Effizienzwert, sondern ein relativer Wert, der sich auf die anderen Universitäten in der Studie bezieht.

DATENSATZ

Der dieser Studie zugrunde liegende Panel-Datensatz bietet mit 70 von insgesamt 76 deutschen öffentlichen Universitäten einen umfassenden Überblick über die Hochschullandschaft in Deutschland, der sich auf den Zeitraum von 2004 bis 2015 bezieht. Der Datensatz enthält Informationen zu den Ausgaben und dem Personal als Input sowie die Anzahl der Absolventen und die Anzahl der hochzitierten Publikationen als Output. Im Gegensatz zu früheren Effizienzstudien deutscher Universitäten (vgl. Gralka, Wohlrabe und Bornmann 2019; Wohlrabe et al. 2019) werden die Universitätskliniken bei den Analysen berücksichtigt.

Die Daten wurden – mit Ausnahme der bibliometrischen Daten – vom Statistischen Bundesamt bezogen. Alle monetären Variablen sind auf das Jahr 2015 deflationiert. Die Informationen zu den Publikationen stammen aus einer hausinternen Datenbank der Max-Planck-Gesellschaft, die auf den Daten des Web of Science (WoS) beruht und von der Max Planck Digital Library (MPDL, München) betrieben wird. In die Auswertungen wurden nur substantielle Publikationen mit den Dokumenttypen »Article« und »Review« einbezogen.

Konform zur klassischen Effizienzanalyse im Hochschulbereich betrachten wir Lehre und Forschung als die primären Aktivitäten von Universitäten. Die Leistungen der Universitäten in der Lehre werden mit der Gesamtzahl der Absolventen aus Bachelor- und Masterstudiengängen (oder äquivalenten Abschlüssen) gemessen. Um den unterschiedlichen Profilen der Universitäten Rechnung zu tragen, wird die Anzahl der Absolventen separat für Naturwissenschaften, Sozial- und Geisteswissenschaften

und Medizin ausgewiesen.¹ Bei der Anzahl von Veröffentlichungen betrachten wir nicht die reine Anzahl, sondern die nach Qualität gewichtete Anzahl. Hierbei betrachten wir die Anzahl der Publikationen, die zu den 10% [$P_{top10\%}$] der am häufigsten zitierten Publikationen in ihrem Fach und Publikationsjahr gehören. Mit der Verwendung dieses fachnormierten Indikators wird in der Studie berücksichtigt, dass man in den wissenschaftlichen Disziplinen unterschiedlich viele Zitate erwarten kann. Den Vergleich von Universitäten, die in mehr als einer Disziplin Forschung betreiben (und publizieren), darf man nur anstellen, wenn man diese Art von normierten Indikatoren verwendet (vgl. Hicks et al. 2015). Waltman et al. (2012) halten $P_{top10\%}$ für den besten Indikator, um institutionelle Vergleiche anzustellen.

Die deskriptiven Statistiken für den vorliegenden Datensatz sind in Tabelle 1 dargestellt (über den gesamten, hier betrachteten Zeitraum). Während im naturwissenschaftlichen Bereich rund 1 000 Studenten pro Jahr ihren Abschluss machen, sind es in den anderen Fächern rund 1 600. In der Medizin beläuft sich die Zahl auf etwa 240. Durchschnittlich rund 150 hochzitierte Publikationen veröffentlicht eine Universität pro Jahr. Die laufenden Ausgaben betragen etwa 260 Mio. Euro jährlich. Rund 3 600 Personen (gemessen in Vollzeitäquivalenten) arbeiten durchschnittlich an einer Universität. Die deskriptiven Statistiken zeigen, dass die Standardabweichung für einen Großteil der Variablen nahe am Mittelwert liegt. Dies deutet auf eine erhebliche Heterogenität der Daten auf institutioneller Ebene hin.

In Abbildung 1 ist die Entwicklung der Variablen über die Zeit dargestellt. Dabei wurden alle Werte in Relation zum Jahr 2004 gesetzt (dessen Werte jeweils 100 bilden), um die Entwicklung bei den Variablen miteinander vergleichen zu können. Es zeigt sich, dass die Anzahl der Absolventen (in den Natur- sowie Geistes- und Sozialwissenschaften) und die Anzahl der hochzitierten Publikationen deutlich angestiegen sind. Die Medizin bildet hier eine Ausnahme:

¹ Die Naturwissenschaften umfassen neben den klassischen Naturwissenschaften, wie die Physik, die Bereiche Mathematik, Agrarwissenschaften, Forstwissenschaften und Ingenieurwesen. Sozial- und geisteswissenschaftliche Fächer sind aus den Bereichen Kunst, Wirtschaft, Recht, Sport und Kultur. Die Medizin umfasst die Human- und Gesundheitswissenschaften sowie die Veterinärmedizin.

Tab. 1

Deskriptive Statistik

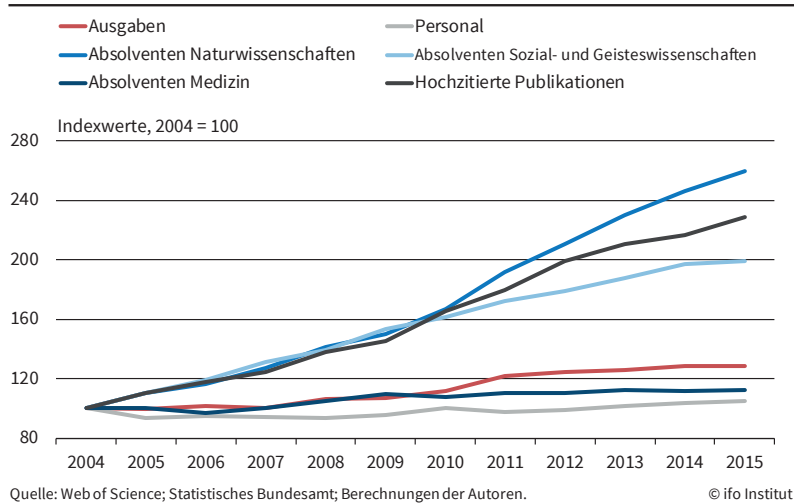
Variable	Mittelwert	Std. Abweichung	Min	Max
Inputs				
Ausgaben ^a	257	183	18	870
Personal ^b	3 589	3 232	179	17 702
Outputs				
Absolventen Naturwissenschaften ^c	1 049	935	0	7 766
Absolventen Sozial- und Geisteswissenschaften ^c	1 580	1 109	0	6 072
Absolventen Medizin ^c	238	297	0	1 361
Hochzitierte Publikationen ^c	157	166	0	950

^a In Mio. Euro. ^b In Vollzeitäquivalenten. ^c In absoluten Werten.

Quelle: Web of Science; Statistisches Bundesamt; Berechnung der Autoren.

Abb. 1

Entwicklung der Input- und Output-Variablen über die Zeit



Bei ihr war der Anstieg eher moderat. Anders als die Output-Variablen zeigen die Input-Variablen deutlich weniger Dynamik. Die Ausgaben sind zwischen 2004 und 2015 nur um rund 28 Prozentpunkte angestiegen; das Personal sogar nur um knapp 5 Prozentpunkte.

ERGEBNISSE

In der klassischen DEA stehen sich die Input- und Output-Variablen aus dem gleichen Jahr gegenüber. Daher beruhen die folgenden Ergebnisse auf einer DEA, bei der die In- und Outputs desselben Jahres eingehen. Insbesondere mit Blick auf die Graduierten stellt sich jedoch die Frage, ob nicht die Ausgaben von mehreren Jahren der Anzahl der Graduierten eines Jahres gegenübergestellt werden sollten. Innerhalb der wissenschaftlichen Literatur wird dies nicht einheitlich gehandhabt (vgl. Aksnes et al. 2016). Deshalb wurde eine zusätzliche Sensitivitätsanalyse durchgeführt, bei der die Inputs mit Hilfe eines Zweijahresdurchschnitts geglättet wurden. Die Ergebnisse sind denen hier präsentierten Ergebnissen sehr ähnlich (Korrelation nach Pearson 0,97). Ein Grund hierfür ist sicherlich die kontinuierlich graduelle Entwicklung der Inputs über die Jahre, d.h. zum Beispiel, dass die Ausgaben von einem Jahr zum nächsten weder stark steigen noch fallen. Die Korrelation der Inputs zwischen den Jahren ist entsprechend hoch. Entsprechend sind keine großen Änderungen bei den Effizienzwerten zu erwarten.

Mit Hilfe der Input- und Output-Variablen wird für jede Universität und jedes Jahr ein Effizienzwert zwischen 0 und 1 berechnet. In Abbildung 2 ist die durchschnittliche Effizienz sowie der jeweilig größte und kleinste Effizienzwert im Zeitverlauf dargestellt. Bei der DEA gibt es immer mindestens eine Universität, die den höchsten Effizienzwert von 1 hat. Die Durchschnittseffizienz über den gesamten Analysezeitraum ist 0,87. Dies bedeutet, dass bei gegebenem Input eine durchschnittliche Universität in Deutschland ihren Output um 13% erhöhen könnte (theoretisch).

Dieser Wert ist ähnlich zu dem Wert, den Wohlrabe et al. (2019) mit einem anderen Datensatz deutscher Universitäten berechnet haben. Wie Abbildung 2 zeigt, ändert sich der durchschnittliche Effizienzwert über die Zeit nur wenig und deutet daher auf Konstanz hin. Die minimalen Effizienzwerte schwanken um einen Wert von 0,5; es ist kein Trend sichtbar.

Abbildung 3 zeigt die Entwicklung der maximal effizienten Universitäten (mit einem Effizienzwert von 1) über die Jahre. Durchschnittlich 21 Universitäten von insgesamt 70 sind pro Jahr maximal effizient. Das Maximum wurde im Jahr 2009 mit 26 Universitäten und das Minimum im Jahr 2015 mit 17 erreicht. Auch wenn die Abbildungen 2 und 3 zeigen, dass die durchschnittlichen Ergebnisse über die Zeit relativ konstant sind, stellt sich die Frage, wie stark die Effizienzwerte der einzelnen Universitäten im Zeitablauf schwanken. Die Antwort wird in Abbildung 4 gegeben. Dort wird sowohl die

Abb. 2

Entwicklung der Effizienzwerte über die Zeit

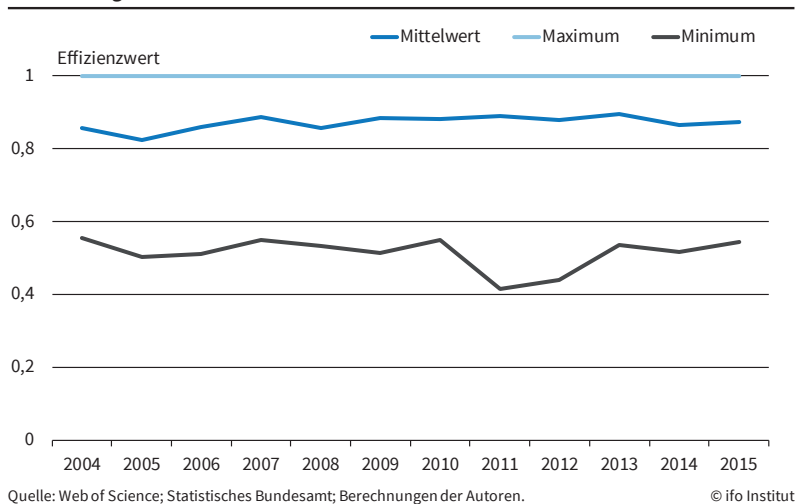
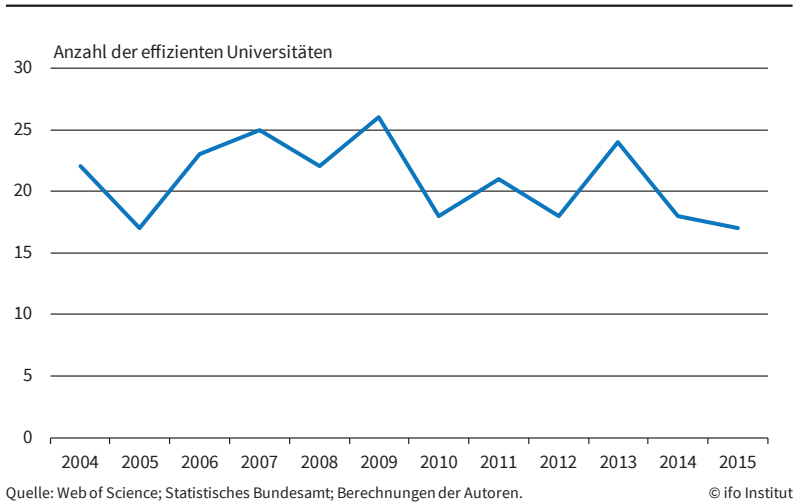


Abb. 3

Anzahl der effizienten Universitäten über die Zeit

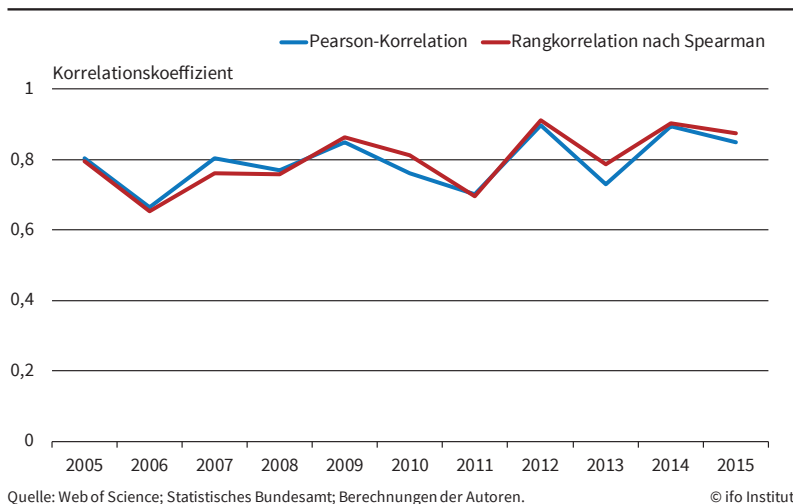


Pearson-Korrelation der Effizienzwerte, als auch die Rangkorrelation nach Spearman dargestellt. Bei der Rangkorrelation werden die Effizienzwerte der Universitäten in Rangpositionen transformiert. Es werden jeweils paarweise Korrelationen zwischen den Jahren berechnet, d.h. dass der Wert 2005 die Korrelation der Werte zwischen 2004 und 2005 darstellt. Abbildung 4 zeigt, dass die Korrelation zwischen den Jahren sehr hoch und im Zeitablauf sogar gestiegen ist. Dabei macht es keinen Unterschied, ob die Korrelation nach Pearson oder Spearman betrachtet wird. Daraus können wir schließen, dass sich die individuellen Effizienzwerte der Universitäten im Zeitablauf nur graduell ändern. Über den gesamten Zeitraum betrachtet liegt die durchschnittliche Differenz zwischen dem höchsten und niedrigsten Effizienzwert einer Universität bei niedrigen 0,23 Effizienzpunkten.

Den Einblick von Abbildung 3 vertiefend, sind in Tabelle 2 die 20 besten Universitäten gemäß ihrer durchschnittlichen Effizienz über die

Abb. 4

Korrelation der Effizienzwerte im Zeitablauf



Zeit dargestellt.² Zusätzlich wird jeweils der (bzw. die) höchste und niedrigste Effizienzwert (Rangposition) aufgelistet. Fünf Universitäten (TU München, LMU München, U Flensburg, U Heidelberg und U Lübeck) waren in jedem Jahr effizient. Die anderen 15 Universitäten waren es in mindestens einem Jahr. Generell gilt dies für 47 von den insgesamt 70 untersuchten Universitäten. Die durchschnittliche Effizienz der Universitäten in Tabelle 2 ist mit mindestens 0,93 sehr hoch.

Die angegebenen Rangpositionen in der Tabelle sind mit Vorsicht zu interpretieren: Auch schlechte Rangplatzierungen weisen sehr gute Effizienzwerte auf. Dies ist damit zu erklären, dass die effizienten Universitäten, d.h. Universitäten mit einem Effizienzwert in Höhe von 1,0, immer den Rang 1 einnehmen. Da, wie in Abbildung 3 gezeigt wird, in jedem Jahr sehr viele Universitäten effizient sind, verschlechtern sich die Rangpositionen der Universitäten, obwohl ihr Effizienzwert nur knapp unterhalb von 1,0 liegt. Deshalb sind die Effizienzwerte aussagekräftiger als die Rangpositionen. Tabelle 2 zeigt zudem, dass institutionelle Effizienz nicht zwangsläufig mit der Größe einer Universität korreliert. So liegt die Korrelation zwischen Ausgaben und Effizienzwerten nur bei etwa 0,2.

FAZIT

Auch wenn deutsche Universitäten in nahezu allen internationalen Universitätsrankings vertreten sind, erhält man auf Basis der Rankings keine Informationen über die tatsächliche Effizienz der Universitäten. Bei den Rankings werden vor allem Output-Indikatoren verwendet. In der vorliegenden Studie haben wir uns deshalb mit der Effizienz der Universitäten in Deutschland beschäftigt. Wie die Ergebnisse zeigen, hat sich die Effizienz der Universitäten über die Zeit wenig geändert. Auch wenn es nur fünf Universitäten gibt, die über den gesamten betrachteten Zeitraum durchgängig effizient sind,

² Die komplette Liste der evaluierten Universitäten findet sich in Galka, Wohlrabe und Bornmann (2019).

Tab. 2

Die 20 effizientesten Universitäten über den gesamten Zeitraum

Rang	Universität	Effizienzwert			Rang		
		Durchschnitt	Minimum	Maximum	Durchschnitt	Minimum	Maximum
1	TU München	1,000	1,000	1,000	1	1	1
1	U Flensburg	1,000	1,000	1,000	1	1	1
1	U Heidelberg	1,000	1,000	1,000	1	1	1
1	U Lübeck	1,000	1,000	1,000	1	1	1
1	LMU München	1,000	1,000	1,000	1	1	1
6	U Koblenz-Landau	1,000	0,998	1,000	3	1	23
7	FU Berlin	1,000	0,994	1,000	3	1	21
8	U Köln	0,997	0,962	1,000	3	1	25
9	U Augsburg	0,992	0,907	1,000	6	1	36
10	U Freiburg i.Br.	0,991	0,915	1,000	9	1	31
11	Humboldt-Universität zu Berlin	0,974	0,703	1,000	7	1	52
12	TU Dresden	0,969	0,879	1,000	14	1	38
13	U Hamburg	0,966	0,841	1,000	16	1	49
14	U Leipzig	0,964	0,849	1,000	13	1	33
15	Bauhaus-U Weimar	0,958	0,805	1,000	16	1	54
16	U Bochum	0,953	0,708	1,000	13	1	56
17	U Karlsruhe	0,950	0,876	1,000	16	1	40
18	U Münster	0,949	0,793	1,000	19	1	40
19	U Ulm	0,946	0,792	1,000	22	1	48
20	U Potsdam	0,938	0,873	1,000	28	1	44

Quelle: Web of Science; Statistisches Bundesamt; Berechnungen der Autoren.

bewegen sich die Effizienzwerte der restlichen Universitäten generell auf einem vergleichbaren Niveau. Mit diesem Ergebnis können wir ein Fazit bestätigen, das bereits häufig zur Position deutscher Universitäten in internationalen Universitätsrankings gezogen wurde: Die deutsche Universitätslandschaft ist eher durch Homogenität als durch Heterogenität gekennzeichnet. In der deutschen Universitätslandschaft existieren nicht die »Leuchttürme der Forschung«, wie sie beispielsweise mit der Harvard University in den USA und der University of Cambridge in Großbritannien vorliegen.

Mit der Exzellenzinitiative hat man in Deutschland den Versuch unternommen, die Homogenität in der Hochschullandschaft aufzubrechen und die Heterogenität zu erhöhen. Es war ein erklärtes Ziel der Exzellenzinitiative, einige »Leuchttürme der Forschung« zu schaffen. Unsere Ergebnisse scheinen darauf hinzudeuten, dass dieses Ziel nicht erreicht werden konnte. Dabei muss man allerdings berücksichtigen, dass wir in der vorliegenden Studie nicht nur den hochzitierten Output berücksichtigt haben, sondern auch den entsprechenden Input gegenübergestellt haben. Bornmann und Haunschild (2016) fragen in diesem Zusammenhang zu Recht, ob man mit Effizienzmessungen im Hochschulbereich die richtigen Anreize setzt. Für Spitzenforschung bedarf es Kreativität, und es ist fraglich, ob Kreativität und Effizienz bei einer Forschungseinrichtung in gleichem Maß gegeben sein können: Kann eine Universität, die besonders effizient ihre Mittel einsetzt, genügend Freiräume für kreative Forschung lassen? Wir können davon ausgehen, dass vor allem innovative Forschung zu »Dead-Ends« führt, was sich negativ auf die Effizienz auswirken würde.

Jede Forschung ist mit mindestens einer Einschränkung versehen – auch unsere. Bei der Interpretation unserer Ergebnisse sollte beachtet werden (wie wir es oben bereits beschrieben haben), dass die Effizienzwerte ein relatives Maß darstellen, die sich aus den gegebenen Daten ergeben haben. Die Ergebnisse zeigen also nicht, ob eine Universität tatsächlich effizient ist, sondern ob sie effizienter ist als andere.

LITERATUR

Agasisti, T., G. Catalano, P. Landoni und R. Verganti (2012), »Evaluating the performance of academic departments: an analysis of research-related output efficiency«, *Research Evaluation* 21, 2–14.

Agasisti, T. und S. Gralka. (2019), »The transient and persistent efficiency of Italian and German universities: A stochastic frontier analysis«, *Applied Economics*, 51(46), 5012–5030.

Agasisti, T. und C. Pohl (2012), »Comparing German and Italian Public Universities: Convergence or Divergence in the Higher Education Landscape?«, *Managerial and Decision Economics* 33, 71–85.

Aksnes, D. W., G. Sivertsen, T. N. van Leeuwen und K. K. Wendt (2016), »Measuring the productivity of national R&D systems: Challenges in cross-national comparisons of R&D input and publication output indicators«, *Science and public policy* 44, 246–258.

Banker, R. D., A. Charnes und W. Wager Cooper (1984), »Some models for estimating technical and scale inefficiencies in data envelopment analysis«, *Management science* 30, 1078–1092.

Bonaccorsi, A., C. Daraio und L. Simar (2006), »Advanced indicators of productivity of universities: An application of robust nonparametric methods to Italian data«, *Scientometrics* 66, 389–410.

Bornmann, L. und R. Haunschild (2016), »Efficiency of research performance and the glass researcher«, *Journal of Informetrics* 10(2), 652–654.

Bornmann, L., M. Stefaner, F. de Moya Anegón, und R. Mutz (2014), »What is the effect of country-specific characteristics on the research performance of scientific institutions? Using multi-level statistical models to rank and map universities and research-focused institutions worldwide«, *Journal of Informetrics* 8(3), 581–593.

Charnes, A., W. W. Cooper und E. Rhodes (1978), »Measuring the efficiency of decision making units«, *European Journal of Operational Research* 2, 429–444.

- De Witte, K. und L. López-Torres (2017), »Efficiency in education: a review of literature and a way forward«, *Journal of the Operational Research Society* 68, 339–363.
- Gawellek, B. und M. Sunder (2016), »The German excellence initiative and efficiency change among universities, 2001–2011«, Working Paper, Universität Leipzig 142.
- Gnewuch, M. und K. Wohlrabe (2018), »Super-efficiency of education institutions: an application to economics departments«, *Education Economics* 26, 610–623.
- Gralka, S. (2018a), »Persistent inefficiency in the higher education sector: evidence from Germany«, *Education Economics* 26, 373–392.
- Gralka, S. (2018b), »Stochastic Frontier Analysis in Higher Education: A Systematic Review«, CEPIE Working Paper 5/18.
- Gralka, S., K. Wohlrabe und L. Bornmann (2019), »How to Measure Research Efficiency in Higher Education? Research Grants vs. Publication Output«, *Journal of Higher Education Policy and Management* 41, 322–341.
- Hazelkorn, E. (2015), *Rankings and the reshaping of higher education: The battle for world-class excellence*, Palgrave Macmillan, New York.
- Hicks, D., P. Wouters, L. Waltman, S. de Rijcke und I. Rafols (2015), »Bibliometrics: the Leiden Manifesto for research metrics«, *Nature News* 520, 429–431.
- Johnes, G. und A. Schwarzenberger (2011), »Differences in cost structure and the evaluation of efficiency: the case of German universities«, *Education Economics* 19, 487–499.
- Kempkes, G. und C. Pohl (2010), »The efficiency of German universities—some evidence from nonparametric and parametric methods«, *Applied Economics* 42, 2063–2079.
- Olivares, M. und H. Wetzel (2014), »Competing in the Higher Education Market: Empirical Evidence for Economies of Scale and Scope in German Higher Education Institutions«, *CESifo Economic Studies* 60, 653–680.
- Rhaimi, M. (2017), »Measurement and determinants of academic research efficiency: a systematic review of the evidence«, *Scientometrics* 110, 581–615.
- Thursby, J. G. (2000), »What do we say about ourselves and what does it mean? Yet another look at economics department research«, *Journal of Economic Literature* 38, 383–404.
- Waltman, Ludo et al. (2012), »The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation«, *Journal of the American society for information science and technology* 63, 2419–2432.
- Warning, S. (2007), *The economic analysis of universities: strategic groups and positioning*, Elgar, Cheltenham.
- Wohlrabe, K., L. Bornmann, S. Gralka und F. de Moya Anegón (2019), »Wie effizient forschen Universitäten in Deutschland, deren Zukunftskonzepte im Rahmen der Exzellenzinitiative ausgezeichnet wurden?«, *Zeitschrift für Evaluation* 18(1), 9–27.