

DRITTMITTEL ALS PERFORMANZINDIKATOR DER WISSENSCHAFTLICHEN FORSCHUNG

Zum Einfluss von Rahmenbedingungen auf Forschungsleistung*

Dorothea Jansen, Andreas Wald, Karola Franke**
Ulrich Schmoch und Torben Schubert***

Zusammenfassung: Ressourcen innerhalb und zwischen Universitäten und in der außeruniversitären Forschung werden angesichts der Finanzknappheit im Wissenschaftssystem zunehmend in Abhängigkeit von Forschungsleistungen verteilt. Hiermit bekommen Evaluationsverfahren und die Messung von Forschungsleistungen einen neuen Stellenwert. Drittmiteileinkommen als relativ einfach zu erhebender Performanzindikator spielt in allen neuen leistungsbasierten Verteilungsverfahren eine große Rolle. Ziel dieses Beitrags ist es, die Aussagekraft von Drittmitteln für die Bewertung von Forschungsleistungen zu untersuchen. Auf der Grundlage von quantitativen und qualitativen Daten zu Forschungsgruppen aus drei sehr unterschiedlichen Forschungsfeldern wird der Zusammenhang von Strukturindikatoren, Drittmiteileinkommen als weiterem Inputfaktor und dem Output der wissenschaftlichen Produktion betrachtet. Hierbei zeigt sich, dass der Einsatz von Drittmitteln als Performanzindikator nur unterhalb eines von disziplinspezifischen Produktionsbedingungen bestimmten Schwellenwertes sinnvoll ist. Ferner ergibt die weitergehende Analyse, dass Forschungsperformanz mehrdimensional ist, Publikationen also nicht der einzig sinnvolle Indikator für Performanz sein können. Schließlich zeigt sich, dass die Bedingungen wissenschaftlicher Produktion zwischen den Fachgebieten so unterschiedlich sind, dass Performanzvergleiche nur innerhalb homogener Forschungsfelder sinnvoll einsetzbar sind. Abschließend werden die forschungspolitischen Implikationen für Forschungsfinanzierung und Forschungsevaluation diskutiert sowie weiterer Forschungsbedarf aufgezeigt.

I. Einleitung

Wissenschaftliche Forschung ist ein Prozess, in dessen Verlauf aus unterschiedlichen Inputfaktoren mittels verschiedener Transformationen Beiträge zur Schaffung neuer wissenschaftlicher Erkenntnisse als Output entstehen. Der Input besteht aus Humanres-

* Die diesem Aufsatz zugrunde liegenden Forschungsarbeiten wurden im Rahmen der von der DFG geförderten Forschergruppe „Governance der Forschung“ (<http://www.foev-speyer.de/governance/>) in den Teilprojekten „Netzwerkstrategien und Netzwerkfähigkeiten von Forschungsgruppen“ (JA 548/4-2) und „Performanzindikatoren“ (SCHM 1719/1-2) durchgeführt.

** Projekt „Netzwerkstrategie und Netzwerkfähigkeit von Forschungsgruppen“ am Forschungsinstitut für öffentliche Verwaltung Speyer.

*** Projekt „Performanzindikatoren für Forschungseinrichtungen, insbesondere Forschungsgruppen“ am Fraunhofer Institut für System- und Innovationsforschung Karlsruhe.

sourcen sowie materiellen Ressourcen, wie beispielsweise Laborausstattungen. Der Output wissenschaftlicher Produktion – die Forschungsleistungen – werden durch wissenschaftliche Kommunikation, in erster Linie Publikationen, aber auch in der Form von Humankapital oder Artefakten sichtbar (Wagner-Döbler 2005). Bewertungen von Forschungsleistungen nach Quantität und Qualität führen zu Aussagen über Forschungsperformanz.

Ressourcen werden angesichts der Mittelknappheit im Wissenschaftssystem zunehmend in Abhängigkeit von Forschungsleistungen und deren Bewertung vergeben. Knappe Mittel sollen so effizient eingesetzt und leistungssteigernde Anreize für Wissenschaftler geschaffen werden (vgl. Göbelbecker 2005; Jaeger 2005; Minssen und Wilkesmann 2003). Hochschulleitungen und Dekane ebenso wie die Leiter von Forschungsinstituten sind vor dem Hintergrund der zunehmenden Autonomie der Universitäten und außeruniversitären Forschungseinrichtungen auf steuerungsrelevante Informationen über die Forschungsleistungen ihrer Fachbereiche und Abteilungen angewiesen.

Voraussetzung für eine leistungsorientierte Mittelvergabe ist eine möglichst exakte und objektive Erfassung und Bewertung der Forschungsleistungen. Während in anderen europäischen Staaten eine breite, regelmäßige Evaluation von wissenschaftlichen Forschungsleistungen längst etabliert ist, befindet sich dieser Prozess in Deutschland noch in einem frühen Stadium. Er gewinnt aber, wie die Rahmenbedingungen des Pakts für Forschung zeigen, erheblich an Dynamik (Cave et al. 1991; Daniel 2001; Liefner 2001; Kehm et al. 2006). Welche Indikatoren für die Messung von Forschungsleistungen herangezogen werden können, ist umstritten (Hornbostel 1997; Larédo 2001). Besondere Probleme bereiten sowohl die Mehrdimensionalität des Forschungsoutputs als auch die oft sehr voraussetzungsvolle und aufwändige Messung. Wichtige Wissenschaftsinstitutionen in Deutschland haben sich daher in letzter Zeit mit den Möglichkeiten zur Messung von Forschungsleistung befasst. Dabei wird dem Drittmitteleinkommen als Indikator eine hohe Bedeutung zugemessen. So veröffentlicht die Deutsche Forschungsgemeinschaft (DFG) neuerdings ein sogenanntes Förder-Ranking (DFG 2003 und 2006). Der Wissenschaftsrat plant alle Fächer der deutschen universitären und außeruniversitären Forschung einem regelmäßigen Rating zu unterziehen (Wissenschaftsrat 2004). In ähnlicher Weise wird Drittmitteln als Indikator für Forschungsleistungen bei der leistungsbezogenen Budgetierung an deutschen Universitäten die größte Relevanz zugemessen (Jaeger 2006: 32).

Das Ziel des vorliegenden Beitrags ist es, die Aussagekraft von Drittmitteln für die Evaluation von Forschungsleistung zu untersuchen. Dabei werden explizit disziplinäre Spezifika und institutionelle Rahmenbedingungen berücksichtigt. Drei sehr unterschiedliche Forschungsfelder werden mit qualitativen und quantitativen Verfahren untersucht. Die Untersuchungseinheiten sind Forschungsgruppen in der Astrophysik, Nanotechnologie und Mikroökonomie. Die betrachteten Fälle repräsentieren alle Säulen des öffentlich finanzierten deutschen Forschungssystems. Damit können – anders als etwa im institutionen- und fächervergleichend angelegten Forschungsranking des CHE (Berghoff et al. 2005) oder dem Förderranking der DFG (2003, 2006) – die Verursachungszusammenhänge zwischen Forschungsbedingungen, Drittmitteleinkommen und Forschungsoutput auf der Mikroebene des Forschungsprozesses für universitäre und au-

ßeruniversitäre Forschungsgruppen untersucht werden und auch qualitative Daten in die Analyse eingehen.

Im folgenden Abschnitt werden die Möglichkeiten der Leistungsmessung in der Forschung und die Rolle von Drittmitteln im Forschungsprozess aus theoretischer Perspektive diskutiert sowie ein Überblick über die Ergebnisse bereits vorliegender empirischer Studien vermittelt. Im dritten Abschnitt werden zunächst die inhaltlichen und strukturellen Besonderheiten der drei untersuchten Felder, das Design der Untersuchung und die Datengrundlage vorgestellt. Der vierte Abschnitt ist der Darstellung der Ergebnisse gewidmet. Zunächst wird auf die disziplinären Unterschiede in der wissenschaftlichen Produktion eingegangen, bevor die quantitativen Struktur- und Leistungsdaten einer multivariaten Analyse unterzogen werden. Für die sich ergebenden Cluster von Forschungsgruppen werden qualitative Daten zur Charakterisierung der typischen Produktionsbedingungen und des Leistungsprofils präsentiert. Abschließend werden Folgen für die Forschung auf dem Gebiet der Wissenschaftsindiktorik diskutiert und Implikationen für die Forschungspolitik aufgezeigt.

II. Performanzindikatoren für die Forschung und Stellenwert von Drittmitteln

In diesem Kapitel wird der Begriff der Forschungsperformanz und die Messbarkeit von Forschung diskutiert. Ferner wird ein Überblick über die Forschung zu den Erklärungsgrößen von Forschungsperformanz gegeben mit einem besonderen Blick auf die Rolle von Drittmitteln. Des Weiteren wird auf die Probleme eingegangen, die im Zusammenhang mit der Institutionalisierung von Verfahren der Performanzmessung in der Forschung auftreten.

1. Forschungsperformanz: Begriff und Messbarkeit

Forschungsleistung lässt sich – anders als klassische Unternehmenskennzahlen – nicht in monetären Größen messen (Jäger 2005). Öffentliche Forschungsorganisationen haben zusätzlich die Schwierigkeit der Mehrdimensionalität ihrer Ziele. Darüber hinaus ist es umstritten, wie sich insbesondere die Qualität wissenschaftlicher Leistung beurteilen lässt, und welche Indikatoren zur Messung der Forschungsleistungen herangezogen werden sollen. Dissens besteht nicht nur hinsichtlich der Vergleichbarkeit zwischen verschiedenen Disziplinen, sondern auch intradisziplinär (Raaijmakers 2004; Hicks 2004). So wird das etablierte Peer Review Verfahren ebenso kritisiert wie bibliometrische Methoden (Bornmann und Daniel 2003: 218–221; Hornbostel 1997: 197ff., 283ff.; Kuhlmann und Heinze 2004a/b; Neidhardt 1988; Zuckerman und Merton 1971).

Das traditionelle Verfahren der Bewertung von Forschungsleistungen, der Peer Review, ist vielfach wegen mangelnder Objektivierbarkeit, kognitivem und disziplinärem Partikularismus und der Konzentration von positiven Bewertungen auf bestimmte etablierte Schulen und Universitäten kritisiert worden (Kruytbosch 1989; Travis und Collins 1991). Kognitiver und institutioneller Partikularismus sind insofern problematisch, als die Chancen für Kreativität von der Offenheit von Organisationen und Sozialsystemen abhängen.

men für Grenzgänger abhängen. Marginalität und Unkonformismus sind häufig eine Voraussetzung für Kreativität (Amabile et al. 1994; Dogan 1999; Williams und Yang 1999). Pioniere, die die Grenzen eines Gebietes überschreiten und Neuland entdecken sowie Hybrid-Akteure, die bisher nicht verbundene Felder und Methoden miteinander verbinden, brauchen flexible Organisationen und flexible Finanzierungsformen, um sich zu entfalten.

Die Kritik an der mangelnden Objektivität des Peer Review führte zur Entwicklung und Nutzung von bibliometrischen Verfahren. Diese Verfahren unterstellen, dass Publikationen der wichtigste sichtbare – und damit messbare – Output wissenschaftlicher Forschung sind (Wouters 1997). Allerdings leiden bibliometrische Verfahren daran, dass sie auf der Basis der vorhandenen Datenbanken nur einen kleinen und spezifischen Ausschnitt von Outputindikatoren betrachten können – Publikationen in ausgewählten, internationalen Journals, Kopublikationen sowie Zitationen. Studien zu den Struktur- und Outputindikatoren von Forschungseinheiten zeigen, dass es organisations- und disziplinspezifische Outputmuster außerhalb des Beobachtungsraums bibliometrischer Indikatoren gibt, etwa eine hohe Bedeutung von Monographien, Editionen und Kommentaren in den sozial- und geisteswissenschaftlichen Fächern oder von Outputformen wie Prototypen, wissenschaftliche Beratung und Technologietransfer in der außeruniversitären oder industriellen Forschung (vgl. Cole 1979; Larédo 2001). Vergleiche zwischen Disziplinen und Forschungsgruppen mit unterschiedlichen Forschungsorientierungen anhand eines einheitlichen Indikatorensatzes stellen sich damit als prekär dar. Ähnliche Probleme ergeben sich durch die unterschiedlichen Publikationsstrategien und die unterschiedliche Journal Coverage bei internationalen Vergleichen innerhalb einer Disziplin, insbesondere bei stark national ausgerichteten Fächern (Schmoch 2004).¹

Problematisch ist ferner die Differenzierung von Produktivität und Qualität von Publikationen. Zitationsraten und Publikationsraten korrelieren zwar tendenziell (Hornbostel 1997: 253ff.), ihr Zusammenhang ist aber so wenig systematisch, dass beide Aspekte separat zu analysieren sind. Selbst zwischen Qualität und Reputation ist keine eindeutige Beziehung feststellbar (Arora et al. 1998; Dewett und Denisi 2004; Makino 1998). Zwar gelten Qualitätsmessungen wie z.B. Zitationen als aussagekräftigeres Maß für Forschungsperformanz. Ihr Einsatz ist jedoch nur mit einem gewissen Zeitverzug zur Publikation möglich. Das notwendige Zitatzeitfenster wird insbesondere bei dem für Kreativität interessanten Bereich der Zitationen zwischen verschiedenen Disziplinen sehr groß (Glänzel et al. 2003; Rinia et al. 2001).

Sowohl Peer Review als auch bibliometrische Indikatoren bilden Handlungen bzw. Bewertungen anderer Wissenschaftler ab und sind insofern sozial und raumzeitlich gebunden. Einzelne Indikatoren bilden nur Teilausschnitte der zu messenden Größe „Beitrag zur Schaffung neuer wissenschaftlicher Erkenntnisse“ ab. Das Ergebnis biblio-

1 Vgl. hierzu zum Beispiel die Kritik, die Gerhards (2002) Analyse der „zwei getrennten Welten“ der Reputation in der deutschen Soziologie gefunden hat, sowohl mit Bezug auf die Mehrdimensionalität von Forschungsleistungen (Publikationen, Organisation von Forschung/Funktionen in Fachgesellschaften und Tätigkeit als Fachgutachter) als auch mit Bezug auf den Befund des geringen Zusammenhangs zwischen nationalen und internationalen Performanzprofilen sowie der problematischen Journal Coverage (Allmendinger 2002; Rammert 2002).

metrischer Recherchen muss daher von geeigneten Experten selbst wieder validiert und eingeordnet werden, was zur Entwicklung des sogenannten Informed Peer Review geführt hat (Daniel 2001; Weingart 2005).

Ein weiteres Problem der Bewertung der unterschiedlichen Leistungsprofile ist der systemische Charakter der Leistungsfähigkeit von Forschungssystemen der Wissenschaft. Ein funktionsfähiges Forschungssystem (auf nationaler oder europäischer Ebene) muss verschiedene Leistungsprofile auf der Ebene von Forschungsorganisationen und Forschungsgruppen aufweisen. Exzellente Forschung ist nur möglich, wenn sie auf Routineforschung mit eher inkrementellem Erkenntniszuwachs aufsetzen kann. Das Forschungssystem braucht eine funktionierende Kommunikationsinfrastruktur, die durch Zeitschriften, Konferenzen, Begutachtungsverfahren, Wissens- und Technologietransferaktivitäten gewährleistet wird. Forschung ist des Weiteren ohne den wissenschaftlichen Nachwuchs nicht möglich, dessen Förderung und Betreuung betrieben werden muss (Stichweh 1994).

Zusammenfassend lässt sich sagen, dass die Messung von Forschungsperformanz sich nicht auf bibliometrische Methoden beschränken kann. Verschiedene Leistungsdimensionen, z.B. auch die Beiträge zur Kommunikationsinfrastruktur des Forschungssystems oder die Förderung des wissenschaftlichen Nachwuchs, müssen in mehrdimensionalen Performanzprofilen abgebildet werden.

2. Determinanten von Forschungsperformanz

Forschung wird im modernen Wissenschaftssystem nicht mehr von Einzelnen betrieben, sondern von Teams oder Forschungsgruppen, die in Organisationen und in nationale und das internationale Wissenschaftssystem eingebettet sind. Diese Gruppen sind in ihrer Leistungsfähigkeit von diesen Kontexten mit geprägt (Bland et al. 2005).

Teamperformanz ist abhängig von Teamvariablen und von Umgebungsvariablen. Zu den Umgebungsvariablen gehört die Einbettung in spezifische Forschungsorganisationen mit unterschiedlichen Forschungsprofilen, die Zugehörigkeit zu Disziplinen und Subdisziplinen mit verschiedenen Produktionslogiken und Wissensdynamiken, sowie die Einbettung in nationale Forschungssysteme und deren Finanzierungsrahmen und Wettbewerbsbedingungen (Adams et al. 2005; Gulbrandson 2004; Hollingsworth 2002; Martin et al. 2004; Whitley 2003). Zu den Teamvariablen zählen die individuellen Merkmale der Teammitglieder, die Merkmale des Teams wie Größe und (disziplinäre) Heterogenität, sowie die Merkmale der Team-Prozesse (Hollingsworth 2002; Stewart 2006; Stock 2004).

Eine aktuelle Studie von Carayol und Matt (2004) zeigt auf der Basis von 88 Instituten der Universität Straßburg, dass Forschungsperformanz (Publikationen) von einer Vielzahl von Faktoren abhängt. Größe und zusätzliches Finanzierungsaufkommen durch die Forschungsförderung von CNRS und INSEAD sowie ein hoher Anteil von Mitarbeitern ohne Lehraufgaben kennzeichnen ein Cluster der großen Forschungslabore, das nur mittlere Performanzwerte erreicht. Institute im Spitzencluster sind dagegen wenig größer als das Cluster mit den kleinsten Einheiten, haben viele in Lehre und Forschung tätige Professoren, viele Doktoranden und einen hohen Anteil an Postdocs

und insbesondere internationalen Postdoc-Positionen.² Auch andere Studien kommen zu dem Ergebnis, dass insbesondere an hochkomplexen Themen arbeitende Gruppen eher vergleichsweise klein sind (Stewart 2006: 45). Die notwendigen Ressourcen und Anregungen können Forschungsgruppen offenbar auch aus der Einbettung in hochreputierte Forschungsorganisationen beziehen.

Detaillierte Studien über den Zusammenhang von Drittmiteileinkommen und Forschungsperformanz sind bisher selten. In einer älteren Studie von Stolte-Heiskanen (1979) konnte insgesamt kein Zusammenhang von Ressourcenausstattung und Performanz festgestellt werden. In der Studie von Hornbostel (1997: 213f.) lassen sich Korrelationen in der Größenordnung von 0,7 bis 0,8 für den Zusammenhang zwischen Publikationsanzahl und Drittmiteileinkommen des Fachbereichs und je Professor ermitteln. Allerdings ist bei einem erhöhten Input an Sach- und Humankapital prinzipiell auch mit einer höheren Zahl an Publikationen zu rechnen. Drittmittel müssen insofern sowohl als Inputfaktor betrachtet werden, der den zukünftigen Forschungsoutput fördert, als auch als Outputfaktor, der das Ergebnis eines Peer Review Prozesses darstellt. Moed et al. (1998) stellen schließlich für flämische Universitäten bei der Drittmittelvergabe eine Konzentration auf wenige Einrichtungen fest und konstatieren speziell für Einrichtungen mit hohem Drittmittelwachstum eine abnehmende Publikationsproduktivität. Sie warnen im Ergebnis vor einer zu starken Ausdehnung der Drittmittelfinanzierung zu Lasten der Grundfinanzierung.

Auf der Ebene von Organisations- und Systemeffekten argumentieren Whitley (2003) und mit Bezug auf Drittmittelmärkte Geuna (1997, 1999) und Laudel (2005). Als entscheidende Faktoren für die Leistungsfähigkeit nationaler Forschungssysteme qualifiziert Whitley (2003) deren Wettbewerbsintensität und gleichzeitige Offenheit. Starke Prestigehierarchien zwischen den Organisationen und jede Begrenzung des Wettbewerbs wirken sich negativ auf die Offenheit und Pluralität von und in Forschungsorganisationen aus und erschweren den Wechsel von Forschungslinien und Forschungskarrieren. Geuna (1997, 1999; vgl. auch Laudel 2005) problematisiert die Tendenz zu hoher Konzentration von fachspezifischen Drittmittelmärkten infolge von Informationsdefiziten und Vergabestrategien der Drittmittelgeber. Neue Bewerber könnten sich hohen Zugangsbarrieren gegenüber sehen. Insbesondere bestehe die Gefahr, dass risikoreiche, aber innovative Projekte nicht ausreichend Unterstützung fänden.

Das zentrale Argument für die Nutzung von Drittmitteln als Performanzindikator besteht darin, dass diese in der Regel im Rahmen eines Gutachterverfahrens vergeben werden und damit eine Bewertung der Forschungsqualität der jeweiligen Forschungseinheit erfolgt (Hornbostel 2001: 524; Wissenschaftsrat 2005). Dem steht allerdings eine Vielzahl von Einwänden gegenüber. Hornbostel führt als zentrales Problem an, dass Drittmittel oft nicht für Forschung, sondern für den Aufbau von Infrastruktur verwendet würden (Hornbostel 2001: 525). Laudel (2005) führt als weitere Einschränkung auf, dass nicht in allen Fällen in ausreichender Weise Förderorganisationen zur Verfügung stünden, was sich insbesondere auf den Vergleich zwischen verschiedenen Ländern und Fachbereichen auswirke. Der gewichtigste Einwand ist, dass ein höherer

² Die Autoren berücksichtigen bei ihrer Untersuchung allerdings nicht den Einfluss disziplinärer Unterschiede, weshalb die Aussagefähigkeit der Ergebnisse stark eingeschränkt ist.

Umfang von Drittmitteln selbstverständlich auch zu einer höheren Zahl von Publikationen führt. Sofern beide Indikatoren in eine Evaluation eingehen, führt dies notwendig zu einer Erhöhung des Gewichts der Drittmittel bei der Leistungsbewertung. Vor diesem Hintergrund kann es in dieser Analyse nicht um den nahe liegenden Zusammenhang von Drittmitteln und Anzahl der Publikationen gehen. Vielmehr ist vor allem die Frage zu beantworten, wie sich Drittmittel auf den relativ zum Input gemessenen Forschungsoutput einer Forschungseinheit auswirken.

Trotz der weit reichenden Akzeptanz der Drittmittel als Leistungsindikator (auch unter den Forschern selbst, vgl. Franke et al. 2006), ist die empirische Basis für die Begründung dieses politischen Steuerungsinstruments bisher sehr dünn. Weitgehend ungeklärt ist, ob der vermutete positive Zusammenhang zwischen Drittmitteln und Forschungsperformanz grundsätzlich besteht. Aus einem produktionstechnischen Kalkül ließen sich durchaus auch andere Verläufe des Zusammenhangs, z.B. abnehmende Erträge und ggf. sogar ein Überwiegen der mit der Drittmittelinwerbung verbundenen Kosten über die Erträge erwarten. Darüber hinaus fehlt es an Wissen über die Abhängigkeit des Zusammenhangs zwischen Drittmitteln und Performanz von weiteren Rahmenbedingungen, wie den institutionellen oder disziplinären Besonderheiten eines Forschungsfeldes. Fach- und institutionenübergreifende einheitliche Steuerungsinstrumente führen in diesem Fall zwar möglicherweise zur verstärkten Einwerbung von Drittmitteln, aber nicht notwendig zu einem größeren und/oder qualitativ besseren Strom von neuen wissenschaftlichen Erkenntnissen.

3. Probleme der Institutionalisierung von Performanzmessung

Obwohl aufgrund der oben referierten Literatur Zweifel an der Reliabilität und Validität von publikations- und drittmittelbasierten Messungen von Forschungsperformanz angebracht sind, werden in vielen standardisierten Mittelvergabeverfahren solche Indikatoren als Leistungskriterium herangezogen (Lezczensky und Orr 2004). Dabei kommt der Höhe der eingeworbenen Drittmittel eine besondere Bedeutung zu (vgl. Berghoff et al. 2005; DFG 2003; 2006; Larédo und Mustar 2000). Ein Grund dafür ist die einfache Messbarkeit dieses Indikators, auf dem viele Allokationsschemata basieren (Franke et. al 2006; vgl. Laudel 2005). Ein zentraler Grund für die Nutzung von Drittmitteln als Performanzindikator ist schließlich darin zu sehen, dass die Grundfinanzierung der Hochschulen in vielen Fächern stagniert oder sinkt und damit ein erhebliches Interesse an der Einwerbung von Drittmitteln besteht.³

Indikatoren, die als Leistungsindikatoren verwendet werden, entwickeln eine durchaus problematische Dynamik und Anreizstruktur. Die Institutionalisierung von Leistungsindikatoren verändert, ähnlich wie reaktive Erhebungsverfahren, das Verhalten der zu Beurteilenden allein durch ihr Vorhandensein (Meyer 1994). So ist z.B. denkbar, dass durch die Einführung des Drittmittelindikators Forschungsgruppen verstärkt Drittmittel-Akquisition betreiben, unabhängig davon, ob dies der Forschung dienlich

³ Nach den vorliegenden statistischen Angaben ist an den Universitäten der Anteil der über Drittmittel finanzierten Forschung zwischen 1980 und 2001 stetig von 22 auf 38 Prozent angestiegen (BMBF 2004; BMFT 1993; HRK 1996; WR 1993, 2000).

ist. Vermutet wird als Konsequenz etwa eine Fokussierung auf Quantität statt Qualität, wenig risikoreiche Mainstream-Forschung und kurzfristig erreichbare Forschungsziele (Weingart 2005: 125ff.). Möglicherweise kann ein Indikator Forschungsperformanz nur abbilden, wenn er nicht zu Evaluationszwecken herangezogen wird. Diese Überlegung trifft allerdings nicht nur auf Drittmittel, sondern auf fast alle Indikatoren zu (vgl. zu Publikationsindikatoren in Australien Butler 2004; Weingart 2005). In jedem Fall ist der Einfluss von Performanzindikatoren, die zur Grundlage von Finanzierungsschlüsseln gemacht werden, auf das soziale Verhalten der Forscher noch nicht hinreichend untersucht.

III. Untersuchungsgegenstand und Untersuchungsdesign

Wir untersuchen den Zusammenhang von Anteilen drittmittelfinanzierter Forschung und Forschungsperformanz für drei Forschungsfelder mit sehr heterogenen Produktionserfordernissen und unterschiedlichem Ressourcenbedarf: Astrophysik, Nanotechnologie und Mikroökonomie. Die Astrophysik wurde als klassisches naturwissenschaftliches Feld mit starker Ausrichtung auf die Grundlagenforschung ausgewählt. Das zweite Feld, die stärker anwendungsorientierte Nanotechnologie, steht für ein mögliches „mode 2“ Forschungsfeld (Nowotny et al. 2003) mit hoher Interdisziplinarität, hohem Einfluss von Industrie und Anwendern auf die Forschung sowie hoher Forschungsdynamik und Konkurrenz. Um neben den Naturwissenschaften auch die in Ressourcenbedarf und Produktionslogik deutlich anders strukturierten Gesellschaftswissenschaften zu repräsentieren, wurde als drittes Feld die Mikroökonomie ausgewählt (vgl. hierzu ausführlicher Franke et al. 2006).

Aufgrund der Abhängigkeit von Großgeräten und Laborausstattung ist die Angewiesenheit von Forschern auf Drittmittel, sowohl was die Größenordnung als auch was den relativen Anteil im Verhältnis zur Grundfinanzierung angeht, in den beiden naturwissenschaftlich-technischen Feldern sehr viel größer. Die Konzentration auf wenige Drittmittelgeber und Prioritätsprogramme ist dabei in der Astrophysik am größten, während Nanotechnologen über eine größere Bandbreite an Fördermöglichkeiten verfügen. In der Mikroökonomie reicht dagegen häufig die Grundfinanzierung aus, um *state of the art* Forschung zu betreiben.

Das Design der Studie zielt auf eine Analyse der Zusammenhänge zwischen Inputvariablen und Forschungsperformanz auf der Mikroebene, d.h. auf der Ebene von Forschungsgruppen. Nur eine solche Disaggregation der Mechanismen und des Outputs erlaubt eine Einsicht in die Verursachungszusammenhänge (Cherchye und Vanden Abeele 2005; Larédo 2001). Unter einer Forschungsgruppe verstehen wir die kleinste eigenständig Forschung betreibende Einheit in einer Organisation, die eigene Forschungslinien langfristig verfolgt. In Universitäten sind dies in der Regel Lehrstühle oder Institute, in außeruniversitären Einrichtungen (z.B. Max-Planck-Instituten) Abteilungen oder Unterabteilungen. Als Grundgesamtheit definieren wir die in den drei ausgewählten Feldern aktiven Forschungsgruppen in allen Säulen des öffentlich finanzierten deutschen Forschungssystem, d.h. in Universitäten und der außeruniversitären Forschung.

Für jedes Forschungsfeld wurde eine Zufallsstichprobe von je 25 Forschungsgruppen in einem mehrstufigen Verfahren identifiziert (ausführlicher: Franke et al. 2006). Zunächst wurden bibliometrisch die in den ausgewählten Forschungsfeldern in den Jahren 2002 und 2003 aktiven deutschen Forschungsorganisationen identifiziert. Im zweiten Schritt wurden die relevanten Forschungsgruppen im Internet recherchiert. Die so generierten Grundgesamtheiten wurden in Expertengesprächen⁴ validiert, bevor schließlich 25 Fälle in einer nach Publikationsaktivität geschichteten Auswahl gezogen wurden. Die institutionelle und regionale Verteilung in der Stichprobe entspricht weitgehend den Verhältnissen in der Grundgesamtheit. Mit den Leitern der Forschungsgruppen wurden zwischen Juni und November 2004 Leitfadengespräche von 60 bis 90 Minuten Dauer zu den Einflüssen der Reformen im deutschen Wissenschaftssystem auf ihre Forschungsstrategien, die Forschungslinien und die Forschungsnetzwerke geführt (Franke et al. 2006; Jansen 2006; Wald et al. 2006). In diesem Rahmen wurde auch ein standardisierter Fragebogen zu den Input-, Struktur- und Outputindikatoren der jeweiligen Forschungsgruppe ausgefüllt. Das hier eingesetzte Indikatorenset ist Ergebnis einer explorativen Vorstudie von 65 Indikatoren in den untersuchten Forschungsfeldern. Für die ausgewählten Indikatoren zeigte sich eine hohe Übereinstimmung sowohl innerhalb der Fächer als auch zwischen ihnen, bei geringfügigen Differenzen hinsichtlich der Relevanzeinschätzung.

Eine disziplinspezifische Auswertung des Indikatorensets findet sich in *Tabelle 1*. Die Inputfaktoren untergliedern sich in die Teilbereiche Personalstruktur, Strukturdaten zur Forschung und technische Infrastruktur. Auf der Inputseite wurde auch der Anteil der Drittmittel als Arbeitszeitanteil der Forschungsgruppe zum Befragungszeitpunkt (2004) erhoben. Ebenfalls über Arbeitszeitanteile in der Forschung wurde die Bedeutung der verschiedenen ausgerichteten Drittmittelgeber abgefragt (z.B. DFG und Stiftungen für grundlagenorientierte Forschung, Industriedrittmittel für anwendungsorientierte Forschung). Zur Erhebung der Forschungsleistungsprofile wurden neben Publikationen in Zeitschriften auch andere häufig in der Literatur erwähnte Indikatoren, wie beispielsweise die Anzahl der Stipendien, Anzahl der Promotionen, Anzahl der Berufungen oder die Anzahl der Forschungspreise erfasst.

In der Auswertungsphase wurde zusätzlich zu den im Rahmen der Interviews erhobenen Indikatoren für Forschungsleistungen für die einzelnen nach Mitgliedern disaggregierten Forschungsgruppen eine bibliometrische Analyse der Publikationen (Zeitschriftenartikel) und diesbezüglichen Kopublikationen und Zitate durchgeführt (1998–2003), in der Astrophysik und der Nanotechnologie auf der Basis des Science Citation Index, in der Mikroökonomie auf der Basis von Scopus. In *Tabelle 1* sind die Drittmittel den Inputindikatoren zugeordnet, wobei sie als Instrumentalziel letztlich zwischen Input- und Outputindikatoren stehen.

⁴ Für die Astrophysik wurden Experten beim Projektträger des BMBF, dem DESY in Hamburg befragt und für die Nanotechnologie wurden Interviews mit dem VDI in Düsseldorf geführt.

IV. Ergebnisse

In diesem Kapitel werden die zentralen empirischen Ergebnisse in vier thematischen Blöcken dargestellt. Zunächst wird geprüft, wie groß die strukturellen Unterschiede zwischen den Fächern/Disziplinen sind und welche Auswirkungen das auf die empirische Vorgehensweise hat. Ein besonders wichtiges Resultat sind die enormen Größenunterschiede der Einheiten unterschiedlicher Fachrichtungen. Es wird deshalb geprüft, ob diese empirisch nachweisbaren Größenvorteile erklärt werden können (Abschnitt IV.1). Weiterhin wird untersucht, ob ein Zusammenhang zwischen Drittmittelquoten und Publikationsperformanz von Forschungseinheiten festzustellen ist und wie dieser aussieht (Abschnitt IV.2). Da Publikationen nur ein sehr unvollständiger Indikator für die Forschungsaktivität von wissenschaftlichen Einheiten sind, wird anschließend die Analyse auf Zitatquoten erweitert, die Beachtung und Qualität wissenschaftlicher Publikationen messen sollen (Abschnitt IV.3). Die Verkürzung der Forschungsmessung auf die Analyse bibliometrischer Indikatoren, auch unter Einschluss von Zitaten, wird von vielen Wissenschaftlern kritisiert. Um diesen Bedenken nachzugehen, wird in einer weiteren Betrachtung das komplette Outputprofil der Forschungseinheiten analysiert, um charakteristische Typen von Einheiten im Hinblick auf ihren Output zu identifizieren (Abschnitt IV.4).

1. Disziplinäre Unterschiede

Die disziplinären Besonderheiten der untersuchten Felder, die bei der Auswahl unterstellt wurden, sollten sich in einer unterschiedlichen Ausprägung sowohl auf Seiten des Inputs als auch des Outputs der wissenschaftlichen Produktion widerspiegeln. Hierzu sind in *Tabelle 1* die disziplinspezifischen Mittelwerte der erhobenen Variablen und signifikante Mittelwertdifferenzen zusammengestellt. Neben Differenzen zwischen den naturwissenschaftlichen Fächern auf der einen und der Wirtschaftswissenschaft auf der anderen Seite zeigen sich auch deutliche Unterschiede zwischen den beiden naturwissenschaftlichen Fächern. Die Unterschiede zwischen den Feldern fallen zum Teil, z.B. bei der Gruppengröße,⁵ sehr deutlich aus. Ähnlich ausgeprägte Unterschiede lassen sich auch beim Drittmittelanteil, insbesondere dem Anteil von Industriedrittmitteln, sowie den Publikationen je Wissenschaftler, den Kopublikationsquoten mit ausländischen Einrichtungen und bei den Zitaten pro Publikation feststellen.⁶ Insgesamt zeigt die Übersicht in *Tabelle 1*, dass, ohne auf einzelne Indikatoren einzugehen, die Fächerdifferenzen so erheblich sind, dass ein einfacher Vergleich von Forschungsgruppen ohne Berücksichtigung der disziplinären Besonderheiten zu erheblich verzerrten Aussagen führen würde. Dies gilt insbesondere auch für die Drittmittelquote.

Als Indikatoren für die publikationsbezogene Forschungsperformanz (Anzahl, Zita-

⁵ Die durchschnittliche Größe der Gruppe liegt in der Astrophysik bei etwa 15, in der Nanotechnologie bei 19 und bei der Mikroökonomie dagegen bei gerade 4.

⁶ Nach den Überlegungen von Hornbostel (2001: 532) könnte der Anteil der Drittmittelfinanzierung in der Mikroökonomie so niedrig sein, dass dieser als Performanzindikator weniger geeignet ist.

Tabelle 1: Input- und Outputvariablen nach Forschungsfeldern und signifikante Feldunterschiede¹

Variable	Mittelwert			Signifikanzniveaus		
	Astro	Nano	Mikro	Astro/ Nano	Nano/ Mikro	Astro/ Mikro
Wissenschaftler	15.33	18.96	4.21		***	***
Anteil grundfinanzierte Wissenschaftler	0.56	0.38	0.86	***	***	***
Anteil drittmittelfinanzierte Wissenschaftler	0.44	0.57	0.14	*	***	***
Disziplinen je Wissenschaftler	0.16	0.27	0.42	**	***	***
Zeitanteil Forschung	60.42	58.60	40.55		***	***
Zeitanteil Lehre	19.79	16.84	34.09		***	***
Zeitanteil Akquisition	10.25	12.88	7.91		***	*
Zeitanteil Sonstiges	9.58	11.68	17.00			*
Anteil Grundlagenforschung	84.00	62.50	74.35	***		*
Anteil angewandte Forschung	2.40	20.38	25.62	***		***
Anteil Entwicklungsforschung	13.60	17.12	0.00		***	***
Anteil Drittmittelforschung	60.21	71.88	20.21		***	***
Anteil DFG-Mittel	35.71	38.01	14.06		***	***
Anteil öffentliche Drittmittel	22.33	23.06	4.85		***	***
Anteil Industriedrittmittel	0.40	6.69	1.48	***	**	
Anteil sonstige Drittmittel	1.77	4.49	0.00		***	***
Zufriedenheit mit Computern	3.88	3.96	3.91			
Zufriedenheit mit Labor	3.39	4.20	entf.	***	entf.	entf.
Zufriedenheit mit Informationsmöglichkeiten	3.96	3.48	3.39	*		**
Zufriedenheit mit Reismitteln	3.32	3.84	3.00		*	
<i>Folgende Indikatoren je Wissenschaftler:</i>						
Stipendien	0.38	0.17	0.22	**		
Kooperationen Unternehmen	0.00	0.18	0.14	***		***
Gutachten Unternehmen	0.00	0.01	0.10		*	**
Promotionen	0.30	0.33	0.30			
Habilitationen	0.08	0.03	0.10		**	
Berufungen	0.03	0.04	0.13		*	**
Preise	0.03	0.08	0.08	**		
Mitgliedschaften in Beratungsgremien	0.55	0.32	0.46	**		
Herausgeberschaften von Zeitschriften	0.03	0.05	0.10			**
Publikationen	2.68	2.40	0.16		***	***
Konferenzbeiträge	3.78	1.89	2.17	***		**
Kopublikationsquoten	69.76	40.58	16.40	***	***	***
Zitate pro Publikation	8.43	4.87	0.24	**	***	***

¹ Zur Analyse der signifikanten Unterschiede zwischen den Disziplinen werden Mittelwertvergleiche bezüglich der erhobenen Variablen herangezogen. Zur Prüfung auf ihre statistische Verlässlichkeit wurden simulationsbasierte Signifikanztests durchgeführt. Dabei wurden Konfidenzintervalle für die Mittelwertdifferenzen am Computer simuliert. Die Nullhypothese der Gleichheit der Erwartungswerte wurde solange nicht abgelehnt, wie der Nullwert Teil dieses Intervalls war. Diese Tests wurden eingesetzt, weil sie günstigere Eigenschaften in kleinen Stichproben und bei unbekannter Verteilung aufweisen als (standardmäßige) doppelte t-Tests.

Lesehilfe: Die Spalten 2 bis 4 geben die Mittelwerte in den Disziplinen an. Die Spalten 5 bis 7 geben die Signifikanzniveaus der Unterschiede an. So indizieren beispielsweise die drei Sterne in der 6. Spalte der Zeile „Wissenschaftler“, dass die mittlere Gruppengröße in der Nanotechnologie von der in der Mikroökonomie auf mindestens 1 Prozent-Niveau abweicht.

* zwischen 10% und 5%; ** zwischen 5% und 1%; *** weniger als 1 %.

tionsquoten, Kopublikationsquoten) sind in *Tabelle 1* die Ergebnisse der in Kapitel III beschriebenen Datenbankrecherche aufgeführt.

2. Zusammenhang von Publikationen und Drittmitteln

Wie bereits in Abschnitt II angesprochen, sollte idealerweise ein streng monotoner Zusammenhang zwischen dem Instrumentalindikator, hier Drittmittelanteil, und dem Fundamentalindikator (Publikationsperformanz) bestehen. Eine Komplizierung des Problems entsteht hier nun noch dadurch, dass die relative Produktivität von Forschung und die Qualität von Forschungsergebnissen auseinander fallen können (vgl. hierzu die Überlegungen in Abschnitt II). Mit den beiden folgenden Analysen soll geprüft werden, ob sich derartige Zusammenhänge zwischen Drittmitteln und Publikationsperformanzindikatoren finden lassen.

Tabelle 2: Multivariate Analyse von Einflussfaktoren auf den Publikationsoutput*

Abhängige Variable: SUMPUBS

Methode: QML Negativ-Binomiales Zählmodell

Einbezogene Beobachtungen: 69

Variable	Koeffizient	Std.Fehler	z-Statistik	P-Wert
C	-0,39206	0,52290	-0,74978	0,45340
MPG	2,13833	0,62659	3,41264	0,00060
FHG_ANDERE	1,64802	0,36648	4,49689	0,00000
A_DRITT	0,09507	0,01885	5,04440	0,00000
ASTRO*A_DRITT^2	-0,00062	0,00017	-3,56691	0,00040
NANO*A_DRITT^2	-0,00055	0,00016	-3,53598	0,00040
MIKRO*A_DRITT^2	-0,00106	0,00023	-4,61955	0,00000
WISS	0,05463	0,01645	3,32085	0,00090
Restr. Log Likelihood	-742,01260		Log Likelihood	-326,00520
LR Statistik (7 df)	832,01470		LR Index (Pseudo-R ²)	0,56065
P-Wert (LR Stat)	0,00000			

* Zu Einzelheiten der verwendeten Methode siehe den Methodenanhang.

MPG:	Max-Planck-Gesellschaft (Dummy)
AUI:	Anderes außeruniversitäres Institut (Dummy)
A_DRITT:	Drittmittelquote
ASTRO*A_DRITT^2:	Quadrierte Drittmittelquote (für Astrophysiker)
NANO*A_DRITT^2:	Quadrierte Drittmittelquote (für Nanotechnologie)
MIKRO*A_DRITT^2:	Quadrierte Drittmittelquote (für Mikroökonomie)
WISS:	Anzahl der Wissenschaftler

Für diese Analyse wurde die Publikationsperformanz, wie im methodischen Anhang genauer ausgeführt, über die Dimension der Publikationsanzahl gemessen. Die Ergebnisse der Regression finden sich in *Tabelle 2*. In die Analyse geht neben dem Anteil der Drittmittelforschung ein quadrierter Term ein (zur Messung möglicher kurvilinearere Beziehungen), der spezifisch für die drei betrachteten Forschungsfelder definiert wurde. Als Kontrollvariablen wurden die Gruppengröße sowie zwei Dummyvariablen zur insti-

tionellen Zugehörigkeit der Forschungsgruppen (Uni, MPG, sonstige außeruniversitäre Forschung) eingeführt.

Alle Koeffizienten mit Ausnahme der Konstanten erweisen sich mindestens auf dem 1 Prozent-Niveau als signifikant (vgl. p-Werte aus *Tabelle 2*).⁷ Der positive Zusammenhang zwischen der Anzahl der Wissenschaftler und dem Publikationsoutput entspricht den Erwartungen. Der positive Koeffizient für die MPG und die anderen außeruniversitären Einrichtungen zeigt, dass die Zugehörigkeit zu einer außeruniversitären Einrichtung die Publikationsproduktivität signifikant erhöht. Hier wirken sich im Vergleich zu Universitäten vermutlich die geringere Lehrbelastung sowie eine bessere Ressourcenausstattung aus.⁸ Im Vergleich der zusammengefassten außeruniversitären Forschungsgruppen mit der MPG spielt vermutlich deren stärker auf anwendungsorientierte Grundlagenforschung orientiertes Profil eine Rolle. Wichtiger als die inhaltliche Interpretation der Koeffizienten der Variablen zur institutionellen Zugehörigkeit ist in diesem Kontext ihre Kontrollfunktion. Die Effekte, die sich in den Koeffizienten der Drittmittelvariablen abzeichnen, sind keine Pseudo-Effekte, die ihren eigentlichen Ursprung in der Differenzierung zwischen universitärer und außeruniversitärer Zugehörigkeit haben.

Bezüglich der interessierenden Effekte der Drittmittelforschung auf die Publikationsperformanz (hier im Sinne von Produktivität), erweist sich trotz der kleinen Fallzahlen der Anteil der drittmittelfinanzierten Forschung als positiv und hochsignifikant. Dies scheint grundsätzlich für die Eignung von Drittmittelforschung als Performanzindikator zu sprechen. Die quadrierten Effekte sind aber für alle drei Forschungsfelder signifikant negativ. Drittmittel wirken demzufolge zunächst positiv auf den Output, ab einem bestimmten kritischen Niveau aber negativ. Dieses kritische Niveau liegt in allen Fällen über den fächerspezifischen Mittelwerten.⁹ Damit dürften nur bei wenigen Forschungseinheiten die Drittmittelstrukturen negativ auf die Publikationsaktivität wirken. Vielmehr befinden sich die meisten Einheiten auf dem steigenden Ast der Publikationsfunktion und könnten deshalb durch zusätzliche Drittmittel Effizienzgewinne realisieren. Allerdings gibt es einige Einheiten, die diesen kritischen Punkt bereits überschritten haben.

Insgesamt ist festzuhalten, dass steigende Drittmittelanteile an der Forschung ab einem bestimmten Niveau negativ auf die Publikationsproduktivität wirken, wobei es deutliche Indizien dafür gibt, dass dieser Effekt vor allem mit überoptimalen Gruppengrößen zusammenhängt. Konkret bedeutet dies, dass zusätzliche Drittmittel, einen hohen Drittmittelanteil gegeben, die absolute Anzahl der Publikationen sinken lässt, obwohl (wahrscheinlich) die Anzahl der Wissenschaftler steigt. Die dahinter liegenden Gründe dürften sein, dass zusätzliche Drittmittel überoptimales Wachstum der Einheiten induzieren können, so dass stark steigende, intraorganisationale Transaktions- und Akquisitionskosten den positiven Effekt zusätzlicher Wissenschaftler überkompensieren.

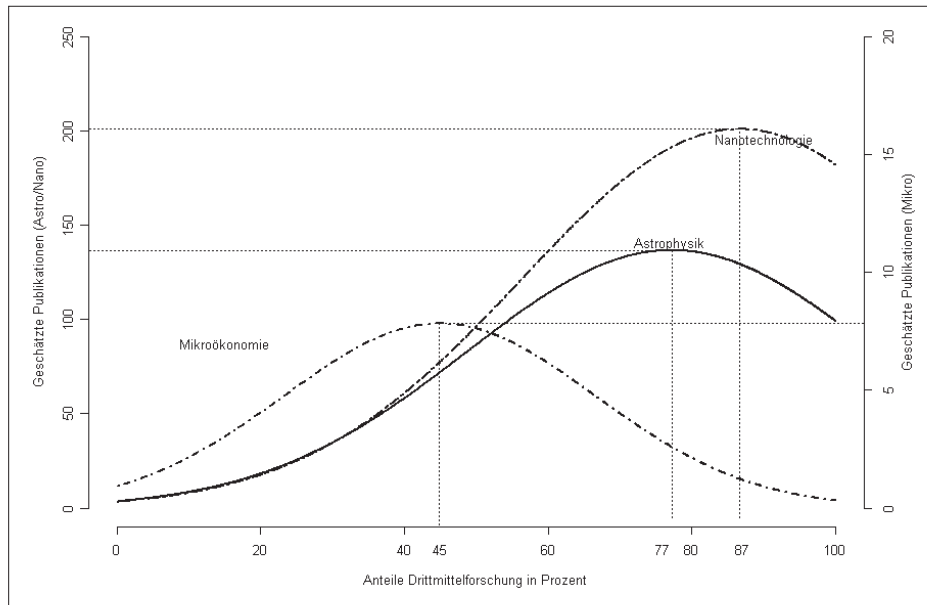
7 Mc-Faddens Pseudo-R² liegt bei 0,56. Die LR-Statistik, die wie eine F-Statistik das restringierte Modell (also nur Konstante) gegen das vollständige Modell schätzt, lehnt ersteres auf nahe 0 Prozent-Niveau ab.

8 Auch die Mitarbeiter und Mitarbeiterinnen von außeruniversitären Einrichtungen sind oft in der Lehre tätig, jedoch meist in einem deutlich geringeren Umfang als Universitätsangehörige.

9 Im Fall der Mikroökonomie liegt der Durchschnitt z.B. bei 20 Prozent gegenüber einem kritischen Wert von etwa 40 Prozent.

Im Ergebnis kann ein monotoner positiver Zusammenhang zwischen Drittmitteln und Publikationsperformanz im Sinne von Produktivität also nicht bestätigt werden. Der Indikator Drittmittel erfüllt nur im Anfangsbereich mit begrenzten Drittmitteln das strenge Erfordernis der Monotonie an einen Performanzindikator (vgl. Abschnitt II).

Abbildung 1: Drittmittelanteile und geschätzte Publikationszahlen für die Forschungsfelder*



* Dass die geschätzte Mittelwertfunktion für die Mikroökonomie bei 100 Prozent Drittmittelquote etwa 0 Publikationen vorhersagt, liegt im Wesentlichen an einigen Einheiten, die sehr hohe Anteile in der Drittmittelforschung haben, aber wenig bis gar nicht publizieren. Bei näherer Betrachtung handelt es sich hier um Einheiten, die sich auf reine Auftragsforschung spezialisiert haben. Der auf den ersten Blick wenig plausible Extremwert erweist sich bei genauer Betrachtung als gar nicht so unrealistisch.

Es lässt sich jedoch ein fester Zusammenhang zwischen Drittmitteln und Publikationsoutput feststellen. Die Höhe des disziplinspezifischen, kritischen Wertes ist prinzipiell bestimmbar, sollte aber im Kontext des beschränkten Datensatzes nicht überinterpretiert werden. In jedem Fall lässt sich mit hoher Signifikanz feststellen, dass diese kritischen Schwellen in der Astrophysik und Nanotechnologie deutlich höher als in der Mikroökonomie liegen, was die obigen Betrachtungen zu den disziplinären Unterschieden hinsichtlich der Gruppengröße aus einem anderen Blickwinkel unterstützt. Die Testergebnisse werden hier aus Platzgründen nicht dargestellt; der Befund ist in *Abbildung 1* illustriert.

3. Typisierung von Forschungsgruppen nach Drittmittel-, Publikations- und Zitatquoten

Im Folgenden wird mit Hilfe einer Clusteranalyse der Zusammenhang von Drittmitteln und Publikationen vertieft betrachtet, wobei als zusätzlicher Indikator für Forschungsperformanz die Variable Zitationen eingeführt wird. Auf diese Weise sollen die Ergebnisse der auf eine Performanzdimension beschränkten Regressionsanalyse zusätzlich validiert werden. Auch sollten unterschiedliche Methoden grundsätzlich zu ähnlichen Ergebnissen kommen. Deshalb werden die Resultate der parametrischen Regressionsanalyse aus dem letzten Abschnitt mit einem nicht-parametrischen Cluster-Ansatz verglichen.

Mithilfe der Ward-Methode konnten drei inhaltlich gut interpretierbare Gruppen identifiziert werden, die in *Tabelle 3* in der Reihenfolge der Anteile für Drittmittelforschung aufgelistet sind. Demnach gibt es bei dem Vergleich zwischen dem Cluster 1 und dem Cluster 2 einen positiven Zusammenhang zwischen der investierten Arbeitszeit in Drittmittelforschung und dem Output, sowohl gemessen in Publikations- als auch Zitatquoten (Performanzquantität und -qualität). *Abbildung 2* illustriert die positive Veränderung bei Publikations- und Zitatquoten zwischen Cluster 1 mit einem unterdurchschnittlichen Drittmittelanteil und Cluster 2 mit einem überdurchschnittlichen Drittmittelanteil. Beim dritten Cluster mit einem sehr hohen Drittmittelanteil sind die Werte sowohl für Publikations- als auch Zitatquoten unterdurchschnittlich. Der Output der Forschungsgruppen geht also im Hinblick auf Produktivität und Qualität zurück. Diese Befunde stehen im Einklang mit den Ergebnissen der multivariaten Analyse. Darüber hinaus erweitern sie diese in Bezug auf die Performanzdimension Qualität/Impact, gemessen durch die Zitatquoten. Auch diese Analyse bestätigt, dass sich die Drittmittelquote nicht unbeschränkt als Indikator für wissenschaftliche Performanz einsetzen lässt.

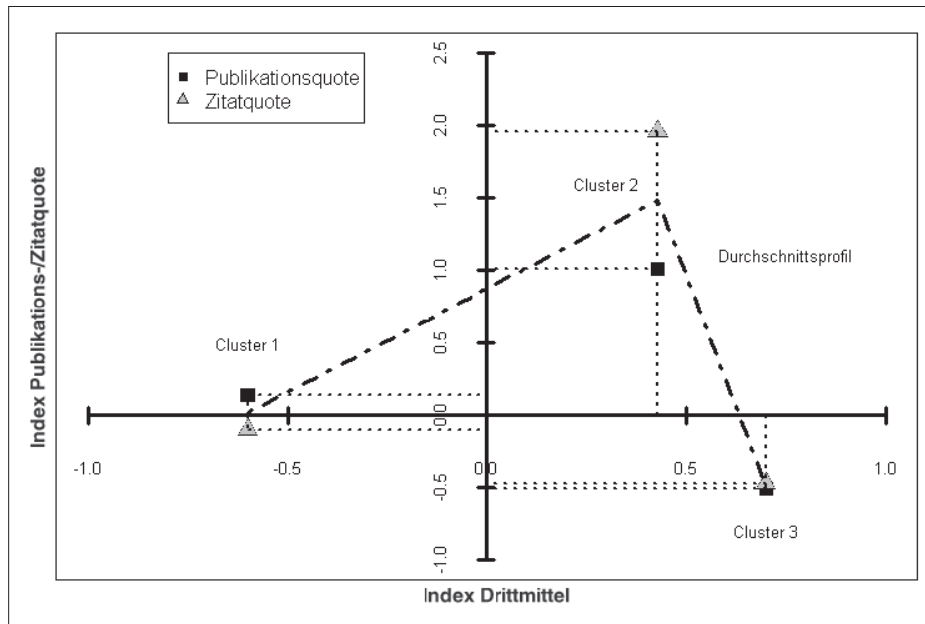
Ein Kritikpunkt an dieser Analyse kann darin gesehen werden, dass Forschungsgruppen aus der Mikroökonomie aufgrund ihres geringen Drittmittelanteils nicht hätten einbezogen werden dürfen. Allerdings ist ihr Anteil in Relation zur Astrophysik und Nanotechnologie gerade im entscheidenden dritten Cluster gering, so dass sich

Tabelle 3: Cluster der Forschungsgruppen nach mittlerem Anteil der Drittmittelforschung, mittlerer Publikationsquote pro Wissenschaftler und mittlerer Zitatquote pro Publikation (disziplinspezifisch standardisierte Werte)*

Gruppe	A-Dritt	Publ.-Quote	Zitatquote	Zahl der Mitglieder
1	-0,60	0,14	-0,10	36
2	0,43	1,01	1,96	8
3	0,70	-0,51	-0,47	26

* In die Clusteranalyse gingen als Variablen die Publikationen pro Wissenschaftler, die Zitate pro Publikation sowie der Anteil der Drittmittelforschung ein. Um die Daten interdisziplinär vergleichbar zu machen, wurden sie jeweils mit dem Intragruppenmittelwert des jeweiligen Faches und der entsprechenden Standardabweichung standardisiert. Die transformierten Werte der Variablen sind größer 0, wenn sie relativ zu ihrer eigenen peer group überdurchschnittlich hohe Werte erreichen. Für einen Mikroökonom wird der Variablenwert also an den Werten aller Mikroökonomien gemessen, nicht aber an denen der Nanotechnologen und Astrophysiker. Sind die Werte kleiner als 0, dann ist das entsprechende Merkmal der Forschungsgruppe unterdurchschnittlich ausgeprägt.

Abbildung 2: Clustermittelwerte für die standardisierten Publikations- und Zitatquoten bezogen auf den standardisierten Anteil der Drittmittelforschung



hieraus keine starke Verzerrung ergibt. Ferner lassen sich zur Validierung der Cluster weitere Informationen aus den Leitfadengesprächen heranziehen, die in einer qualitativen Inhaltsanalyse zu Variablen vercodet worden sind.

Das am stärksten besetzte Cluster 1 ($n=36$), das Gruppen mit unterdurchschnittlichem Drittmittelanteil, leicht überdurchschnittlicher Publikations- und leicht unterdurchschnittlicher Zitatquote beinhaltet, enthält überproportional viele Mikroökonomien. Gleichzeitig sind diese Forschungsgruppen vergleichsweise klein und haben einen geringen Drittmittelbedarf. Im Vergleich mit den beiden anderen Clustern ist Cluster 1 durch den höchsten Anteil außeruniversitärer Institute gekennzeichnet, wobei insbesondere Max-Planck-Institute stark vertreten sind. Dem entspricht ein deutlich geringerer Anteil angewandter Forschung. Für die Forschungsgruppen im ersten Cluster spielt die Drittmitteleinwerbung bei der internen Ressourcenallokation die geringste Rolle. Die Forschungsgruppen in diesem Cluster weisen die kleinsten Forschungskooperationsnetzwerke auf und werden in der Wahl ihrer Kooperationspartner am wenigsten durch externe Anreize, z.B. von Drittmittelgebern, beeinflusst.

Das am schwächsten besetzte Cluster 2 ($n=8$) enthält Gruppen mit einer überdurchschnittlichen Drittmittelquote sowie deutlich überdurchschnittlichen Publikations- und Zitatquoten. Nur eine der Gruppen stammt aus der außeruniversitären Forschung. Drittmittel haben eine deutlich höhere Bedeutung für die Forschungsfähigkeit der Gruppen als im ersten Cluster. Die Gruppen weisen die größten Forschungskooperationsnetzwerke auf, die vergleichsweise stark von externen Anreizen, zum Beispiel den Anforderungen von Drittmittelgebern beeinflusst sind. Es handelt sich um große

Gruppen mit einem höheren Anteil angewandter Forschung und Industriekooperationen. Im Vergleich zu den beiden anderen Clustern bewerten die Forscher in Cluster 2 die Unterstützung durch die Administration der eigenen Institution positiv, sind mit ihrer Ressourcenausstattung zufrieden und halten den Aufwand zur Einwerbung von Drittmitteln für nicht zu hoch. Gegenüber den beiden anderen Clustern ist für die Gruppen im Cluster 2 der Arbeitszeitanteil, der auf die Drittmittelakquisition entfällt, am geringsten.

Cluster 3 (n=26) beinhaltet die Gruppen mit dem höchsten Drittmittelanteil sowie unterdurchschnittlichen Publikations- und Zitatquoten. Bis auf zwei Gruppen stammen alle aus Universitäten. Die Astrophysik ist überproportional, die Mikroökonomik unterproportional vertreten. Vor allem für die Astrophysik lässt sich ein Überwiegen des Struktureffekts der Drittmittel vermuten, da die Mittel überproportional für die Infrastrukturausstattung bzw. den Zugang zu Rechnerkapazitäten und Teleskopen eingesetzt wurden. Diese Investitionen setzen sich möglicherweise erst langfristig in Publikationen um. Die Inspektion dieses Clusters vor dem Hintergrund der qualitativen Daten zeigt darüber hinaus einige weitere Messprobleme auf, die durch Mobilität der leitenden Forscher oder durch ein hohes Drittmittelinkommen für Auftragsforschung – möglicherweise ohne internationale Publikationsmöglichkeiten – entstehen können. Aufschlussreich ist auch die qualitative Analyse der Rahmenbedingungen dieser Forschungsgruppen. Die Gruppen betreiben den relativ höchsten Aufwand für die Akquisition von Drittmitteln (s.o.). Die Abhängigkeit von Drittmittelinwerbungen zur Gewährleistung von Forschungsfähigkeit ist in diesem Cluster am ausgeprägtesten. Drittmittel spielen zudem in der internen Ressourcenallokation im Vergleich zu den anderen Clustern eine größere Rolle. Die Unterstützung durch die eigene Verwaltung wird ebenso wie die eigene Infrastruktur als unzureichend und deutlich negativer als in den anderen Clustern bewertet. Die Forschungspolitik insgesamt ebenso wie die Förderpolitik der Drittmittelgeber werden stark negativ bewertet. Die Situation in Cluster 3 illustriert die schwierigen Bedingungen von Forschungsgruppen, deren Drittmittelquote oberhalb des oben diskutierten optimalen Wertes liegt.

4. Mehrdimensionalität wissenschaftlichen Outputs

Forschungsperformanz ist eine latente Variable, die sich durch bibliometrische Indikatoren nur ausschnittartig beschreiben lässt. In den Interviews mit den Forschungsgruppenleitern wurde daher ein umfangreiches Outputprofil mit insgesamt 13 unterschiedlichen Indikatoren einschließlich weiterer Publikationsindikatoren wie Monographien und Aufsätze in nationalen referierten Zeitschriften erhoben (vgl. Wald 2005 sowie *Tabelle 1* in Abschnitt IV.1). In diesem Abschnitt soll die Mehrdimensionalität des wissenschaftlichen Outputs thematisiert werden. Dieser umfasst neben den durch bibliometrische Kennzahlen repräsentierten Forschungsleistungen, andere Aspekte, die in der Befragung durch weitere Outputindikatoren erfasst wurden. Wie in Abschnitt II ausgeführt, wird von vielen Autoren bezweifelt, dass eine Performanzanalyse ausschließlich auf der Basis bibliometrischer Kennzahlen den besonderen Bedingungen des Wissenschaftsbetriebs ausreichend Rechnung tragen kann. Neben den beiden in den Analysen

im Abschnitt IV.2 und IV.3 untersuchten Performanzdimensionen wurden drei weitere Dimensionen deutlich:

- Aufrechterhaltung der wissenschaftlichen Infrastruktur (Herausgeberschaften, Tätigkeit in wissenschaftliche Beratungsgremien),
- Förderung des wissenschaftlichen Nachwuchses (Habilitationen, Promotionen),¹⁰
- Wissens- und Technologietransfer sowie Kooperationen mit Unternehmen.

Zur weiteren Verdichtung der Daten wurde auf der Basis der Faktorwerte für die fünf Performanzdimensionen eine Clusteranalyse durchgeführt, in der vier Gruppen identifiziert wurden. Ihre Profile bezüglich der Performanzdimensionen zeigt *Abbildung 3*.

Als wesentliches Ergebnis dieser Analyse zeigt sich, dass sich die beiden Gruppen auf der rechten Seite von *Abbildung 3* gut über bibliometrische Indikatoren abbilden lassen. Diese beiden Gruppen, die hier plakativ als „Publizierer“ und „Hochzitierte“ bezeichnet werden, umfassen etwa 40 Prozent des Datensatzes. Die beiden anderen Gruppen sind demgegenüber nicht notwendigerweise weniger leistungsfähig, sondern weisen ein anderes Aktivitätsprofil aus. So zeichnen sich die „Nachwuchsförderer“ über eine erhöhte Aktivität im Bereich des Forschernachwuchses (über den Faktor „Habilitationen/Promotionen“) aus. Die andere Gruppe der „Netzwerker“ (in etwa 30 Prozent) ist neben der Bereitstellung von Organisationsleistung über Herausgeberschaften auch im Bereich wissenschaftlicher Beratungsgremien aktiv. Was die Forschungsleistung betrifft, tendiert diese Gruppe bei der wissenschaftlichen Beachtung (Zitate, Kopublikationen) eher in Richtung Durchschnitt und ist bei den Publikationen schwach, wohingegen sich bei der zweiten Gruppe der „Nachwuchsförderer“ (in etwa 30 Prozent) die Situation genau umgekehrt darstellt.

Trotz der statistischen Grenzen dieser Betrachtung können doch zwei Folgerungen abgeleitet werden. Zunächst einmal besteht eine Tendenz zur funktionalen Differenzierung (Aktivitätsspezialisierung), womit ähnliche Untersuchungen von Larédo und Mustar (2000) im Bereich der Genetik bestätigt werden.¹¹ Das bedeutet letztlich, dass Forschungsgruppen i.d.R. nicht in allen Performanzdimensionen überdurchschnittliche Werte erreichen.

Darüber hinaus fällt auf, dass fast alle Forschungsgruppen aus Cluster 3 (87 Prozent), wie es in Abschnitt IV.3 ermittelt wurde,¹² zu den „Netzwerkern“ und den „Nachwuchsförderern“ gehören. Die Einheiten, die in Abschnitt IV.2 als performanzschwach charakterisiert wurden, übernehmen also andere Funktionen in der Lehre und bei der Organisation der wissenschaftlichen Infrastruktur, also Aufgaben, die für die Funktion der Scientific Community elementar sind.

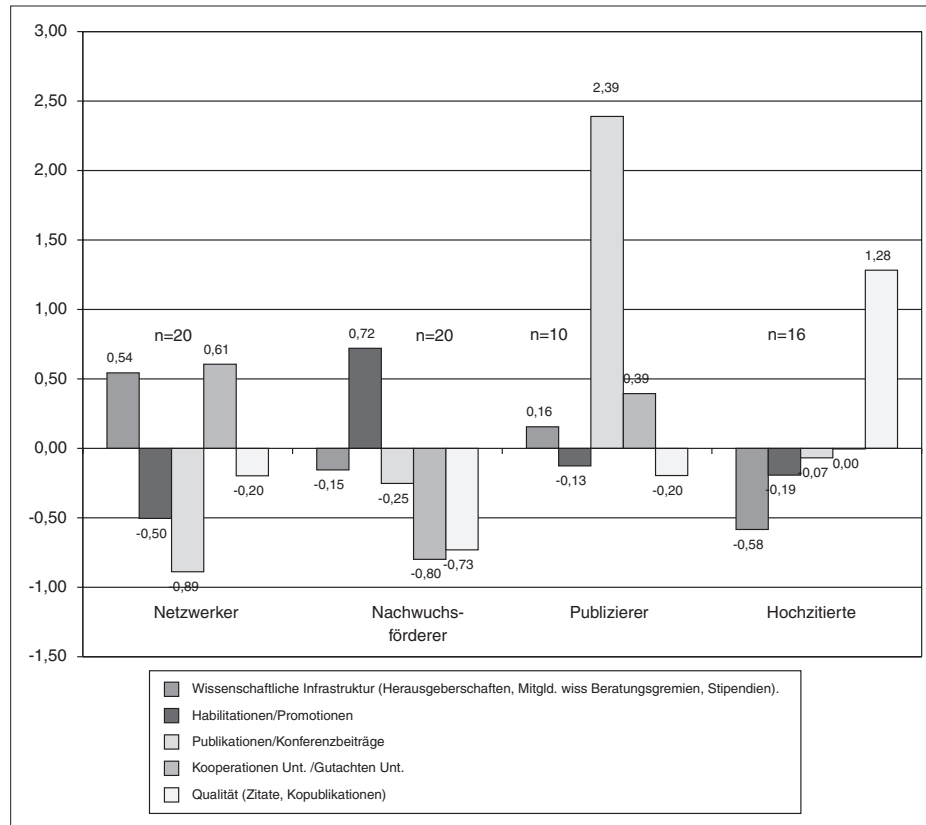
Hinsichtlich der Eignung bibliometrischer Methoden für die Evaluation der Performanz wissenschaftlicher Tätigkeit ist daher zu konstatieren, dass diese zwar ein wichti-

10 Die Förderung des wissenschaftlichen Nachwuchses wird als Element des Forschungsoutputs verstanden, da sie die langfristige Funktionsfähigkeit des Forschungssystems sichert. Förderung (Ausbildung) von Nachwuchswissenschaftlern hat somit Investivcharakter und sollte deshalb (im Gegensatz zur beruflichen Qualifizierung im Studium) nicht als originärer Teil der Lehre verstanden werden.

11 Larédo und Mustar nehmen eine etwas andere Typisierung vor. Vgl. zu unterschiedlichen Outputprofilen auch schon Andrews (1979).

12 Charakterisiert durch hohen Anteil der Drittmittelforschung und unterdurchschnittliche Outputs in Publikationen und Zitationen.

Abbildung 3: Forschungsgruppentypen nach Outputdimensionen*



* Um charakteristische Profile im Outputportfolio der verschiedenen Forschungseinheiten zu verdeutlichen wurde eine Faktorenanalyse der in *Tabelle 1* zusammengestellten Auswahl von Outputindikatoren – mit Ausnahme des Indikators „Forschungspreise“ – durchgeführt. Für diese Analyse wurden die Indikatoren wiederum einer disziplinspezifischen z-Transformation unterzogen. Die Anzahl der Eigenwerte größer als 1 betrug 5. Als Ergebnis konnten fünf gut interpretierbare Faktoren extrahiert werden.

ges Instrument sind, aber für sich allein gestellt nur ein unzureichendes Bild der wissenschaftlich relevanten Leistungen widerspiegeln. Die bibliometrischen Indikatoren sind daher um weitere Indikatoren zu ergänzen, um die Mehrdimensionalität der Performanz in adäquater Weise zu erfassen. Auch bei Forschungseinheiten mit Leistungsprofilen, die nicht auf Publikationen fokussiert sind, ist allerdings ausnahmslos eine Publikationsaktivität festzustellen. Auch für diese Einheiten sind bibliometrische Analysen sinnvoll, soweit sie das niedrigere Durchschnittsniveau der jeweiligen Referenzgruppe in Rechnung stellen. Bibliometrische Analysen sollten daher nicht pauschal, sondern differenziert nach Leistungsprofilen eingesetzt werden.

V. Schlussfolgerungen

Die obige Analyse führt zu dem Ergebnis, dass die Nutzung von Drittmitteln als Performanzindikator für die Forschung nur in einem Bereich mit begrenztem Anteil von Drittmittelforschung sinnvoll ist. Ab einem kritischen Schwellenwert, der in den Natur- und Technikwissenschaften höher als in den Gesellschaftswissenschaften liegt, besteht jedoch der positive Effekt zusätzlicher Drittmittel auf den wissenschaftlichen Output nicht mehr. Dabei sinkt nicht nur die Produktivität gemessen in der Zahl der Publikationen pro Wissenschaftler, sondern auch die absolute Zahl der Publikationen. Das hat nicht nur Konsequenzen für die unkritische Nutzung von Drittmitteln als Performanzindikator, sondern auch wissenschaftspolitische Implikationen. Denn der in Deutschland und anderen Ländern zu beobachtende Trend einer allmählichen Reduzierung der Grundmittel kann offensichtlich nicht durch eine Erhöhung der Drittmittel vollständig kompensiert werden. Es handelt sich vielmehr um ein Phänomen, welches negative Konsequenzen für die wissenschaftliche Performanz haben kann. Dies gilt besonders dann, wenn die Drittmittelvergabe starke Konzentrationstendenzen aufweist. Eine mögliche Folge wäre, dass einige (wenige) Institute deutlich wachsen – eventuell über ihre optimale Gruppengröße hinaus – und viele andere weit unter ihrer mindest-optimalen Größe verharren. Es ist auch zu bedenken, dass sehr hohe Drittmittelanteile die Rahmenbedingungen für die Forschungseinheiten so verändern können, dass auch die Erträge der Forschung selbst davon negativ berührt werden. Die Folgerungen aus beiden Überlegungen sind unterschiedlich. Träfe das erste Argument zu, würden zu große Forschungsgruppen allgemein schlecht abschneiden. Der negative Effekt hoher Drittmittel folgte aus der Tatsache, dass sich in Zeiten knapper Haushalte einige übergroße Gruppen, andererseits viele unterkritische Gruppen herausbilden würden. Hingegen würden gemäß der zweiten Überlegung große Gruppen deshalb schlechter abschneiden, weil sie sich sehr stark über Drittmittel finanzierten. Einige der qualitativen Charakterisierungen zum Cluster 3 der Analyse in Abschnitt IV.2 deuten auch in diese Richtung. Die Folge wäre, dass ein großes Institut mit hohen Drittmittelanteilen seine Leistungsfähigkeit erhöhen könnte, indem es das Finanzierungsportfolio in Richtung stärkerer Grundfinanzierung veränderte. Da die beiden Überlegungen sich nicht gegenseitig ausschließen, werden sie vermutlich parallel eine gewisse Berechtigung haben.

Die Untersuchungen zeigen weiterhin, dass Performanzvergleiche mit Hilfe von Indikatoren nur innerhalb von homogenen Fachgebieten eingesetzt werden sollten, wohingegen Vergleiche zwischen Fachgebieten zu nicht adäquaten Aussagen führen. Die Bedingungen wissenschaftlicher Produktion zwischen Disziplinen sind so unterschiedlich, dass dies in die Betrachtung eingehen muss. Eine Mindestanforderung ist eine Normierung auf der Basis der disziplinären Durchschnittswerte. Darüber hinaus haben die Input- und Outputindikatoren aber auch ein nach Disziplinen unterschiedliches relatives Gewicht, was adäquat in Rechnung zu stellen ist. In dieser Hinsicht waren auf der Basis des vorliegenden Datensatzes nur begrenzte Aussagen möglich. Deshalb ist geplant, den Datensatz erheblich auszuweiten, um tiefer gehende Differenzierungen zwischen Fachgebieten zu ermöglichen.

Die vorliegende Analyse ermöglicht weiterhin, der Frage nach der Aussagefähigkeit

verschiedener Outputindikatoren hinsichtlich der wissenschaftlichen Performanz nachzugehen. Es bestätigt sich dabei, dass, wie in der Literatur oft qualitativ postuliert, Publikationen nicht der einzig sinnvolle Indikator für Performanz sein können. Vielmehr lassen sich typische mehrdimensionale Performanzprofile ermitteln, wobei aber ca. 40 Prozent der untersuchten Forschungsgruppen den Clustern zuzurechnen sind, bei denen die Publikationsaktivität bzw. der Impact (Zitationen, Kopublikationen) im Vordergrund steht. Bei den verschiedenen Profilen gibt es kein geeignetes Verfahren, um diese in eine eindeutige Rangfolge zu bringen bzw. einen optimalen Mix innerhalb eines Fachgebiets zu identifizieren. Sie alle haben in dem komplexen Gefüge der Organisation der Wissenschaft ihre funktionale Berechtigung. Die Frage nach einer adäquaten Relation zwischen den verschiedenen Leistungsprofilen ist letztlich ein wissenschaftspolitisches Problem. Innerhalb der verschiedenen Aktivitätscluster lassen sich allerdings bibliometrische Analysen weiterhin anwenden; problematisch ist die pauschale Anwendung auf alle Forschungsgruppen innerhalb eines Fachgebietes.

Auch wenn es im Sinne der Performanzmessung keine eindeutige Präferenz für ein bestimmtes Leistungsprofil gibt, so lassen sich in negativer Hinsicht Forschungsgruppen bestimmen, die sich in keiner Performanzdimension zumindest im Durchschnittsbereich befinden. Weiterhin sind sinnvolle Vergleiche zwischen Gruppen möglich, die demselben Cluster zugehören.

Auch wenn sich gezeigt hat, dass Drittmittelforschung nur begrenzt als Performanzindikator verwendbar ist, so geben sie gute Hinweise im Hinblick auf die Governancestrukturen, da in der aktuellen Situation die Struktur der Drittmittel wesentlich die Orientierung der Forschungsgruppen beeinflusst. So führt beispielsweise schon ein geringer Anteil an Drittmittelforschung für Unternehmen zu einer stärkeren Anwendungsorientierung der Forschung insgesamt, um nur ein nahe liegendes Ergebnis zu nennen. Eine genauere Analyse dieser Frage sowie eine disziplinspezifische Betrachtung der Zusammenhänge zwischen spezifischen Input- und Outputindikatoren sind wesentliche Gründe für die geplante, oben angesprochene Erweiterung des Datensatzes.

Insgesamt macht die Untersuchung deutlich, dass ein hinreichend differenzierter Datensatz es ermöglicht, eine Reihe von Fragen zu beantworten, die sich aus qualitativen Überlegungen zur Performanz der Wissenschaft ergeben. Gleichzeitig bleiben jedoch viele Punkte offen, so dass die Empfehlung einer Verbindung von quantitativer und qualitativer Analyse im Sinne eines Informed Peer Review bestehen bleibt.

A. Methodenanhang

Als Outputindikator für Forschung wurden Artikel in international referierten Fachzeitschriften herangezogen. Diese sind positive und ganzzahlige Größen, so dass nur Zählmodelle als regressionsanalytische Instrumente geeignet sind. Zunächst wurde ein Poisson-Modell angepasst und ein Test auf die Dispersion durchgeführt. Dieser lehnte das Poisson-Modell auf 2 Prozent-Niveau ab. Mit dem geschätzten Dispersions-Parameter von 0,26, der „overdispersion“ indiziert, wurde das in *Tabelle 2* dargestellte NegBin-Modell geschätzt. Dispersion bezeichnet dabei die unterstellte Form der Heteroskedastizität. Im Poisson-Modell steigt die Varianz proportional mit dem geschätzten Erwartungswert. Für das NegBin-Modell gilt: $V(x_i) = \mu_i (1 + m \mu_i)$, wobei m der Dispersionsparameter ist. Wie man sieht, ist die Varianzstruktur für $m=0$ im NegBin- und im Poisson-Modell identisch. Implementiert wurde die Schätzung als Quasi-Maximum-Likelihood Schätzung mit he-

teroskedastizitätskonsistenten Standardfehlern. Das unterstellte Modell liefert auch dann konsistente Schätzer und Standardfehler, wenn die unterstellte Negativ-Binomial-Verteilung eine unzutreffende Annahme ist.

Als unabhängige Variablen wurden die Institutionentypen (Max-Planck-Gesellschaft, andere außeruniversitäre Einrichtungen, Universitäten), die Drittmittelquoten sowie ihre quadrierten Terme definiert. Zusätzlich wurde als unabhängige Variable die Anzahl der Wissenschaftler eingeführt. Sowohl die Institutionentypen als auch die Wissenschaftler sollen als Kontrollvariablen verstanden werden. Gegen eine solche regressionsbasierte Analyse kann eingewandt werden, dass zwar auf der einen Seite die Publikationen von den Drittmitteln abhängen, aber auf der anderen Seite der Auswahlprozess bei der Drittmittelvergabe vergangene Publikationen berücksichtigt. Die Drittmittelquoten sind dann nicht exogen, sondern simultan mit den Publikationen bestimmt. Durch einen entsprechenden Test wurde in der Tat die Simultaneität der Drittmittelanteile und der Publikationen und damit ihre wechselseitige Abhängigkeit bestätigt. Die Einführung einer entsprechenden Korrektur durch Schätzung eines Mehrgleichungsmodells führte allerdings zu Effekten gleicher Richtung mit etwas geringerer Signifikanz. Dieses Mehrgleichungsmodell wurde mit 3-stage-least-squares (3SLS), einem geeigneten Verfahren, das auf Instrumentenvariablen beruht, geschätzt. Wegen der geringen Unterschiede haben sich die Autoren entschlossen, das einfacher interpretierbare Modell (Tabelle 2) darzustellen.

Literatur

- Adams, James D., Grant C. Black, J. Roger Clemmons und Paula E. Stephan, 2005: Scientific Teams and Institutional Collaborations: Evidence from U.S. Universities, 1981–1999. *Research Policy* 34: 259–285.
- Allmendinger, Jutta, 2002: Eine drei-Welten-Lehre wissenschaftlicher Reputation und ihre Messung. *Soziologie* 31: 56–58.
- Amabile, Teresa M., Regina Conti, Heather Coon, Jeffrey Lazenby und Michael Herron, 1994: The Context of Innovation. The Role of the Work Environment. Waltham, Mass.: Brandeis University.
- Arora, Ashish, Paul A. David und Alfonso Gambardella, 1998: Reputation and Competence in Publicly Funded Science: Estimating the Effects on Research Group Productivity. *Annales D'Economie Et De Statistique* 49–50: 163–198.
- Berghoff, Sonja, Gero Federkeil, Petra Giebisch, Cort-Denis Hachmeister, Detlef Müller-Böling und Meike Siekermann, 2005: Das CHE-Forschungsranking deutscher Universitäten 2004. Gütersloh: CHE-Arbeitspapier Nr. 62.
- Bland, Carole J., Bruce A. Center, Deborah A. Finstad, Kelly R. Risbey und Justin G. Staples, 2005: A Theoretical, Practical, Predictive Model of Faculty and Department Research Productivity. *Academic Medicine* 80: 225–237.
- Bornmann, Lutz, und Hans-Dieter Daniel, 2003: Begutachtungen durch Fachkollegen in der Wissenschaft. Stand der Forschung zur Reliabilität, Fairness und Validität des Peer-Review-Verfahrens. S. 207–225 in: Stefanie Schwarz und Ulrich Teichler (Hg.), *Universität auf dem Prüfstand. Konzepte und Befunde der Hochschulforschung*. Frankfurt a.M./New York: Campus.
- Bundesministerium für Bildung und Forschung (BMBF), 2004: Bundesbericht Forschung. Berlin/Bonn: BMBF.
- Bundesministerium für Forschung und Technologie (BMFT), 1993: Bundesbericht Forschung 1993. Bonn: BMFT.
- Butler, Linda, 2004: What Happens when Funding is Linked to Publications Counts? S. 389–405 in: Henk F. Moed, Wolfgang Glänzel und Ulrich Schmoch (Hg.), *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems*. Dordrecht/Boston/London: Kluwer Academic Publishers.

- Carayol, Nicolas, und Mireille Matt, 2004: Does Research Organization Influence Academic Production? Laboratory Level Evidence from a Large European University. *Research Policy* 33: 1081–1102.
- Cave, Martin, Stephen Hanney und Maurice Kogan, 1991: *The Use of Performance Indicators in Higher Education. A Critical Analysis of Developing Practice*. 2nd edition, London: Kingsley.
- Cherchye, Laurens, und Piet Vanden Abeele, 2005: On Research Efficiency. A Micro-analysis of Dutch University Research in Economics and Business Management. *Research Policy* 34: 495–516.
- Cole, Gerald A., 1979: Classifying Research Units by Patterns of Performance and Influence: a Typology of the Round 1 data. S. 353–404 in: *Frank M. Andrews* (Hg.), *Scientific Productivity*. Cambridge: Cambridge University Press.
- Daniel, Hans-Dieter, 2001: *Wissenschaftsevaluation*. Center for Science and Technology Studies. Bern. CEST 2001/2.
- Deutsche Forschungsgemeinschaft (DFG), 2003a: Förder-Ranking 2003. Institutionen – Regionen – Netzwerke – DFG-Bewilligungen und weitere Basisdaten öffentlich geförderter Forschung. Bonn.
- Deutsche Forschungsgemeinschaft (DFG), 2006: Förder-Ranking 2006. Institutionen, Regionen – Netzwerke. DFG-Bewilligungen und weitere Basisdaten öffentlich geförderter Forschung. Bonn.
- Dewett, Todd, und Angelo S. Denisi, 2004: Exploring Scholarly Reputation: It's More than just Productivity. *Scientometrics* 60: 249–272.
- Dogan, Mattei, 1999: Marginality. *Encyclopedia of Creativity* 2: 179–184.
- Franke, Karola, Andreas Wald und Katinka Bartl, 2006: Die Wirkung von Reformen im deutschen Forschungssystem. Eine Studie in den Feldern Astrophysik, Nanotechnologie und Mikroökonomie. Speyerer Forschungsbericht, Speyer, im Erscheinen.
- Garcia, Clara Eugenia, und Luis Sanz-Menéndez, 2005: Competition for Funding as an Indicator of Research Competitiveness. *Scientometrics* 64: 271–300.
- Gerhards, Jürgen, 2002: Reputation in der deutschen Soziologie – zwei getrennte Welten. *Soziologie* 31: 19–33.
- Geuna, Aldo, 1997: Allocation of Funds and Research Output: The Case of UK Universities. *Revue d'Economie Industrielle* 79: 143–162.
- Geuna, Aldo, 1999: *The Economics of Knowledge Production. Funding and the Structure of University Research*. New Horizons in the Economics of Innovation. Cheltenham, UK/Northampton, MA: Elgar.
- Glänzel, Wolfgang, und Andrés Schubert, 2004: Analysing Scientific Networks through Co-authorship. S. 257–276 in: *Henk F. Moed, Wolfgang Glänzel und Ulrich Schmoch* (Hg.), *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Glänzel, Wolfgang, Balázs Schlemmer und Bart Thijs, 2003: Better Late Than Never? On the Chance to Become Highly Cited only Beyond the Standard Bibliometric Time Horizon. *Scientometrics* 58: 571–586.
- Göbelbecker, Jürgen, 2005: The Role of Publications in the New Programme Oriented Funding of the Hermann von Helmholtz Association of National Research Centres (HGF). *Scientometrics* 62: 173–181.
- Gulbrandson, Magnus, 2004: Accord or Discord? Tensions and Creativity in Research. S. 31–57 in: *Sven Hemlin, Ben Marti und Carl Martin Allwood* (Hg.), *Creative Knowledge Environments: the Influences on Creativity in Research and Innovation*. Cheltenham: Edward Elgar.
- Hicks, Diana, 2004: The Four Literatures of Social Science. S. 473–496 in: *Henk F. Moed, Wolfgang Glänzel und Ulrich Schmoch* (Hg.), *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Hochschulrektorenkonferenz (HRK) (Hg.), 1996: *Zur Finanzierung der Hochschulen. Dokumente zur Hochschulreform 110/1996*. Bonn: Hochschulrektorenkonferenz.
- Hollingsworth, Roger, 2002: *Research Organizations and Major Discoveries in Twentieth-Century Science: A Case of Excellence in Biomedical Research*. Berlin: WZB Discussion Paper P02-003.

- Hornbostel, Stefan*, 1997: Wissenschaftsindikatoren. Bewertungen in der Wissenschaft. Opladen: Westdeutscher Verlag.
- Hornbostel, Stefan*, 2001: Third Party Funding of German Universities. An Indicator of Research Activity? *Scientometrics* 50: 523–537.
- Jaeger, Michael*, 2005: Leistungsbezogene Mittelvergabe und Qualitätssicherung als Elemente der hochschulinternen Steuerung. Hochschul-Informations-System A12/2005, Hannover.
- Jaeger, Michael*, 2006: Leistungsbezogene Budgetierung an deutschen Universitäten. *Wissenschaftsmanagement* 3: 30–36.
- Jansen, Dorothea*, 2006: Research Networks – Origins and Consequences: Preliminary Evidence from a Study of Astrophysics, Nanotechnology and Micro-economics in Germany. In: *Max Albert* (Hg.), *Scientific Competition. Theory and Policy. Conferences on New Political Economy* 24, im Erscheinen.
- Kruytbosch, Carlos E.*, 1989: The Role and Effectiveness of Peer Review. S. 69–85 in: *David Evered* und *Sara Harnett* (Hg.), *The Evaluation of Scientific Research*. Chichester, UK: Wiley.
- Kuhlmann, Stefan*, und *Thomas Heinze*, 2004a: Evaluation von Forschungsleistungen in Deutschland. Erzeuger und Bedarf. Teil I: Konzeptionelle Grundlage. *Wissenschaftsrecht* 37: 53–69.
- Kuhlmann, Stefan*, und *Thomas Heinze*, 2004b: Evaluation von Forschungsleistungen in Deutschland. Erzeuger und Bedarf. Teil II: Produktion und Verwendung evaluativer Information sowie Möglichkeiten ihrer künftigen Organisation. *Wissenschaftsrecht* 37: 125–149.
- Larédo, Philippe*, und *Philippe Mustar*, 2000: Laboratory Activity Profiles: An Exploratory Approach. *Scientometrics* 47: 515–539.
- Larédo, Philippe*, 2001: Benchmarking of RTD Policies in Europe: Research Collectives as an Entry Point for Renewed Comparative Analysis. *Science and Public Policy* 28: 285–294.
- Laudel, Grit*, 2005: Is External Research Funding a Valid Indicator for Research Performance. *Research Evaluation* 14: 27–34.
- Leszczensky, Michael*, und *Dominic Orr*, 2004: Staatliche Hochschulfinanzierung durch indikatoren-gestützte Mittelverteilung. – Dokumentation und Analyse der Verfahren in 11 Bundesländern. Kurzinformation Hochschul-Informations-System A2/2004. Hannover.
- Liefner, Ingo*, 2001: Leistungsorientierte Ressourcensteuerung in Hochschulsystemen. Ein internationaler Vergleich. Berlin: Duncker & Humblot.
- Makino, Junichiro*, 1998: Productivity of Research Groups – Relation between Citation Analysis and Reputation within Research Communities. *Scientometrics* 43: 87–93.
- Merton, Robert K.*, 1968: The Matthew Effect in Science. *Science* 57: 68–72.
- Meyer, Marshall W.*, 1994: Measuring Performance in Economic Organizations. S. 556–578 in: *Neil J. Smelser* und *Richard Swedberg* (Hg.), *The Handbook of Economic Sociology*. Princeton/N.J.: Princeton University Press.
- Minssen, Heiner*, und *Uwe Wilkesmann*, 2003: Lassen sich Hochschulen steuern? *Soziale Welt* 54: 123–144.
- Moed, Henk F.*, 2005: Citation Analysis in Research Evaluation. Berlin u.a.: Springer.
- Moed, Henk F.*, *Marc Luwel*, *Josee A. Houben*, *Eric Spruyt* und *Herman Van Den Berghe*, 1998: The Effects of Changes in the Funding Structure of the Flemish Universities on their Research Capacity, Productivity and Impact during the 1980's and early 1990's. *Scientometrics* 43: 231–255.
- Neidhardt, Friedhelm*, 1988: Selbststeuerung in der Forschungsförderung. Das Gutachterwesen der DFG. Opladen: Westdeutscher Verlag.
- Nowotny, Helga*, *Peter Scott* und *Michael Gibbons*, 2003: Mode-2 revisited. *The New Production of Knowledge*. *Minerva* 41: 179–194.
- Raan van, Anthony F.J.* (Hg.), 1988: *Handbook of Quantitative Studies of Science and Technology*. Amsterdam u.a.: North-Holland.
- Raan van, Anthony F.J.*, 2004: Measuring Science. *Capita Selecta of Current Main Issues*. S. 19–50 in: *Henk F. Moed*, *Wolfgang Glänzel* und *Ulrich Schmoch* (Hg.), *Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems*. Dordrecht/Boston/London: Kluwer Academic Publishers.

- Rammert, Werner, 2002: Die halbierte Reputation – eine grob fahrlässige und unfaire Rechnung! *Soziologie* 31: 53–55.
- Rinia, Ed J., Thed N. Van Leeuwen, Eppo E. W. Bruins, Hendrik G. Van Vuren und Anthony F. J. Van Raan, 2001: Citation Delay in Interdisciplinary Knowledge Exchange. *Scientometrics* 51: 293–309.
- Schmoch, Ulrich, 2004: Leistungsfähigkeit und Strukturen der Wissenschaft im internationalen Vergleich 2004. Studien zum deutschen Innovationssystem Nr. 6-2005. Bundesministerium für Bildung und Forschung. Berlin.
- Stewart, Greg L., 2006: A Meta-Analytic Review of Relationships Between Team Design Features and Team Performance. *Journal of Management* 32: 29–54.
- Stichweh, Rudolf, 1994: *Wissenschaft, Universität, Professionen: Soziologische Analysen*. Frankfurt a.M.: Suhrkamp.
- Stock, Ruth, 2004: Drivers of Team Performance: What Do We Know and What have we Still to Learn? *Schmalenbach Business Review* 56: 274–306.
- Stokes, Terry D., und James A. Hartley, 1989: Coauthorship, Social Structure and Influence within Specialities. *Social Studies of Science* 19: 101–125.
- Stolte-Heiskanen, Veronica, 1979: Externally Determined Resources and the Effectiveness of Research Units. S. 121–153 in: Frank M. Andrews (Hg.), *Scientific Productivity*. Cambridge: Cambridge University Press.
- Travis, G. David L., und Harry M. Collins, 1991: New Light on Old Boys: Cognitive and Institutional Particularism in the Peer Review System. *Science, Technology, & Human Values* 16: 322–342.
- Wagner-Döbler, Roland, 2005: The System of Research and Development Indicators: Entry Points for Information Agents. *Scientometrics* 62: 145–153.
- Wald, Andreas, 2005: Zur Messung von Input und Output wissenschaftlicher Produktion. Daten und Ergebnisse einer Untersuchung auf der Ebene von Forschungsgruppen. *FÖV Discussion Papers* 20, Speyer.
- Wald, Andreas, Karola Franke und Dorothea Jansen, 2006: Governance Reforms and Scientific Production. Evidence from German Astrophysics. In: Dorothea Jansen (Hg.), *New Forms of Governance in Research Organizations – From Disciplinary Theories towards Interfaces and Integration*. Dordrecht: Springer (forthcoming).
- Weingart, Peter, 2005: Impact of Bibliometrics upon the Science System: Inadvertent Consequences? *Scientometrics* 62: 117–131.
- Whitley, R., 2003: Competition and Pluralism in the Public Sciences: the Impact of Institutional Frameworks on the Organisation of Academic Science. *Research Policy* 32: 1015–1029.
- Williams, Wendy M., und Lana T. Yang, 1999: Organizational Creativity. S. 373–391 in: Robert J. Sternberg (Hg.), *Handbook of Creativity*. Cambridge: Cambridge University Press.
- Wissenschaftsrat, 1993: *Drittmittel der Hochschulen 1970 bis 1990*. Köln: Geschäftsstelle des Wissenschaftsrates.
- Wissenschaftsrat, 2000: *Thesen zur künftigen Entwicklung des Wissenschaftssystems in Deutschland*. Köln.
- Wissenschaftsrat, 2004: *Empfehlungen zu Rankings im Wissenschaftssystem. Teil 1: Forschung*. Hamburg.
- Wissenschaftsrat, 2005: *Stellungnahme zu Leistungsfähigkeit, Ressourcen und Größe universitätsmedizinischer Einrichtungen*. Bremen, S. 60.
- Wooldridge, Jeffrey M., 1999: *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass./London: The MIT Press.
- Wouters, Paul, 1997: Citation Cycles and Peer Review Cycles. *Scientometrics* 38: 39–55.
- Zuckerman, Harriet, und Robert K. Merton, 1971: Patterns of Evaluation in Science. Institutionalisation, Structure and Functions of the Referee System. *Minerva* 9: 66–100.

Korrespondenzanschrift: Prof. Dr. Dorothea Jansen, Deutsche Hochschule für Verwaltungswissenschaften Speyer, Lehrstuhl für Soziologie der Organisation, Postfach 1409, 67324 Speyer
E-Mail: Ls_jansen@dhw-speyer.de