

Reputationsdimensionen in den Sozialwissenschaften

Der Preis der Fritz Thyssen Stiftung und
andere wissenschaftliche Leistungsmerkmale¹

Marcus Jansen

In der sozialwissenschaftlichen Forschungsgemeinde wird kaum ein Thema so kontrovers diskutiert wie die Evaluierung ihrer eigenen Leistung. Forschungspolitische Interessen, Paradigmenpluralismus und die Selbstbezüglichkeit der Wissenschaftsevaluation, in der die Eignung wissenschaftlicher Leistungen für die wissenschaftliche Kommunikation durch wissenschaftliche Methoden ermittelt wird (Hirschauer 2004), erschweren eine von allen beteiligten Akteuren getragene Übereinkunft über ein einheitliches Evaluationsverfahren, und infolgedessen trifft nahezu jede Bemühung der Evaluation von Forschungsleistung auf vehemente Kritiker. Alber und Fliegner (2010) unterstreichen, dass aufgrund der Vielschichtigkeit von wissenschaftlichen Leistungen diese nicht anhand eines einzelnen Kriteriums bewertet werden können und weisen auf die Notwendigkeit facettenreicher Evaluationsprozesse hin. Forschungseinrichtungen pflegen die In-

¹ Der Artikel entstammt einer umfangreichen Untersuchung, die unter dem Titel »Die Evaluation deutschsprachiger Sozialwissenschaft im Zeitraum 1981 bis 2006 durch das Prämierungsverfahren des Preises der Fritz Thyssen Stiftung« von der Fritz Thyssen Stiftung gefördert wird, und profitiert erheblich von den Studien der frühen Projektphase, die Jens Alber, Florian Fliegner und Torben Nerlich am WZB durchführten. Heiner Meulemann und Karl-Wilhelm Grümer bin ich für hilfreiche Anmerkungen und Informationen zu Dank verpflichtet.

frastruktur wissenschaftlicher Kommunikation, bilden den wissenschaftlichen Nachwuchs aus, werben Drittmittel ein und publizieren neue Wissensserkenntnisse. Jede dieser Formen wissenschaftlichen Handelns bildet eine Reputationsdimension. Sie »messen [...] jeweils Unterschiedliches und fügen sich dennoch zu einer Welt« (Allmendinger 2002: 56).

Eine bis vor kurzem von der Evaluationsforschung gänzlich unberücksichtigte Form der Konstruktion wissenschaftlicher Exzellenz ist der Preis der Fritz Thyssen Stiftung für deutschsprachige sozialwissenschaftliche Aufsätze, der 1983 erstmals für den Zeitschriftenjahrgang 1981 vergeben worden ist. Dieser Wissenschaftspreis ist Untersuchungsgegenstand eines an der Universität zu Köln in Kooperation mit dem Wissenschaftszentrum Berlin für Sozialforschung durchgeführten und von der Fritz Thyssen Stiftung finanzierten Forschungsprojektes. In der frühen Projektphase konzentrierte sich die Untersuchung ausschließlich auf 94 Aufsätze, die im Prämierungsverfahren zwischen den Jahren 1981 und 2006 zum Zuge kamen (Alber, Fliegner 2010; Alber et al. 2010). Neben der Bestimmung der Merkmale der prämierten Aufsätze und Autoren wurde der Zitationserfolg prämierter Autoren und das Abschneiden im Prämierungsverfahren erfolgreicher Forschungseinrichtungen bei anderen Standortevaluationen untersucht. Da in diesen Untersuchungen nicht nominierte Aufsätze unberücksichtigt blieben, verzerrt die Fokussierung auf absolute Prämierungserfolge die Messung der Erfolge zugunsten großer Forschungsstandorte. Diese haben tendenziell einen höheren Publikationsoutput und sind in der Grundgesamtheit der Artikel der teilnehmenden Zeitschriften stärker vertreten. Folglich ist – ceteris paribus – rein statistisch die Chance einer Nominierung durch die Zeitschriftenherausgeber und einer anschließenden Prämierung durch die Jury für größere Einrichtungen höher als für kleine Einrichtungen.

Mit der nun abgeschlossenen Vollerhebung aller 5.285 am Prämierungsverfahren beteiligten Aufsätze können Aussagen über relative Erfolge getroffen werden und der Zusammenhang zwischen dem Distinktionsmerkmal *Thyssen Preis* und anderen Leistungsmerkmalen der Forschung kann unverzerrt geschätzt werden. Die bisherige Evaluationsforschung bringt im Fach Soziologie eine mehrfache *zwei-Welten-Lehre* wissenschaftlicher Reputation zum Ausdruck. Demnach sind im Bereich der Sozialwissenschaft nicht nur Wissenschaftsindikatoren wie Zahl der Publikationen und Anzahl der bewilligten Forschungsanträge, sondern auch andere Leistungsmerkmale kaum miteinander korreliert. Die vorliegende Studie untersucht, ob eine Reputierung mit dem Leistungsmerkmal *Thyssen Preis* mit anderen wissenschaft-

lichen Leistungen einhergeht oder dieser Wissenschaftspreis eine eigene Welt im System der wissenschaftlichen Exzellenzzuschreibung ist.

Die Untersuchung nähert sich der Forschungsfrage mit einer theoretischen Diskussion, in der das Distinktionsmerkmal *Thyssen Preis* in die Typologie der Leistungsmerkmale und Wissenschaftsindikatoren eingebettet und Hypothesen über den Zusammenhang zwischen den verschiedenen Leistungsmerkmalen formuliert werden. Der Beschreibung des Datensatzes und der methodischen Vorgehensweise der Datenanalyse folgt die empirische Untersuchung. Auf der Mikroebene wird hierzu der Rezeptionserfolg einzelner Aufsätze in Bezug zu deren Erfolg im Prämierungsverfahren gesetzt. Auf der Makroebene der Forschungseinrichtungen wird der Erfolg im Prämierungsverfahren mit den Ergebnissen der Hochschulrankings des Centrums für Hochschulentwicklung (CHE) und den Ergebnissen des Forschungsratings Soziologie des Wissenschaftsrats verglichen. Die zentralen Ergebnisse werden im letzten Abschnitt zusammengefasst.

Theorie

Die Heterogenität von wissenschaftlichen Leistungsmerkmalen hat mehrere Ursachen. Münch (2007: 274, 295f.) weist darauf hin, dass in den Geistes- und Sozialwissenschaften bei der Bewertung von Drittmittelanträgen und Publikationen heterogene Standards existieren. Gutachter von Zeitschriften und Forschungsanträge können zum Beispiel denselben paradigmatischen Ansatz oder dieselbe thematische Ausrichtung einer Forschungsarbeit unterschiedlich bewerten. Zudem ist die Drittmittelinwerbung mit einem erheblichen zeitlichen Aufwand verbunden und diese Ressourcenbindung wirkt sich in derselben Periode negativ auf die Forschungsaktivitäten und die Publikation von Forschungsergebnissen aus.

Darüber hinaus werden Korrelationen falsch geschätzt, wenn die Wissenschaftsindikatoren nur ungenau erfasst werden. Regelmäßig werden Zweifel an der Validität von in bibliometrischen Studien gemessenen Publikationsindikatoren geäußert. Kritisiert werden unter anderem fehlende Gewichtungen hinsichtlich der Autorenanzahl, Seitenlänge des Beitrags, Typus des Forschungsbeitrags (Replik, Rezension, Aufsatz) oder der Auswahl der berücksichtigten Publikationsmedien. So wurde zum Beispiel Gerhards' Studie (2002) stark kritisiert (Allmendinger 2002; Endrweit

2002; Rammert 2002), da er für die Generierung des individuellen Publikationsoutputs nur zwei Zeitschriften auswertete und gar keine Gewichtung vornahm. Die Replikationsstudie von Schulz-Schaeffer (2002) kann die Schwächen der Untersuchung von Gerhards auch nur zum Teil beseitigen.

Unter den Evaluationsstudien der institutionellen Ebene wird die Aussagekraft der Hochschulrankings des Centrums für Hochschulentwicklung ebenfalls regelmäßig angezweifelt (Ursprung 2003: 178). Mit der ausschließlichen Berücksichtigung von Veröffentlichungen in begutachteten internationalen Zeitschriften bedient sich das letzte Ranking deutscher Hochschulen im Bereich der Soziologie (Berghoff et al. 2008) wieder einer fragwürdigen Methode. Die Fokussierung auf internationale Zeitschriftenartikel vernachlässigt die diversifizierte Publikationskultur in der deutschsprachigen Soziologie. Zwischen den Jahren 2001 und 2005 sind nur 15,6% der Publikationen deutschsprachiger Soziologinnen und Soziologen im nicht-deutschsprachigen Ausland erschienen. Mehr als die Hälfte aller Publikationen waren Beiträge zu Sammelwerken oder Herausgeberschaften solcher Sammelwerke, und 7% der Veröffentlichungen waren Monographien (Wissenschaftsrat 2008: 36). Mit dem hohen Anteil deutschsprachiger Veröffentlichungen bedient die Soziologie nach Meinung des Wissenschaftsrates die starke Nachfrage seitens Politik, Verbänden, Wirtschaft und Öffentlichkeit nach Analysen nationaler und regionaler Problemstellungen. Dieser Leistungsbereich wird in der Evaluierung des CHE gänzlich ausgespart. Neben der inhaltlichen Kritik ergibt sich zusätzlich das Problem, dass das Ranking aufgrund der niedrigen Publikationszahlen besonders anfällig ist gegenüber Änderungen des Publikationsverhaltens der Forscher. Schon geringe Variationen des internationalen Publikationsoutputs der einzelnen Universitäten hätten erhebliche Verschiebungen im Ranking zur Folge. Durch einen Anstieg der Veröffentlichungen um jährlich drei Publikationen könnte zum Beispiel die Universität Konstanz von Platz 24 auf Platz 7 im Publikationsranking des Fachs Soziologie steigen, wenn der Output der übrigen Universitäten konstant bliebe.

Ungeachtet der bisher erwähnten Ursachen für das Auseinanderklaffen verschiedener Wissenschaftsindikatoren ist die Stimmigkeit der Annahme der Korrelation unterschiedlicher Leistungsmerkmale grundsätzlich zu hinterfragen (Allmendinger 2002). Die Leistungserbringung in der Wissenschaft erfolgt in unterschiedlichen Bereichen und daher existieren mehrere Reputationsdimensionen. Aufgrund der Multidimensionalität wissenschaftlicher Leistungen müssen Indikatoren, die diese Leistungen zu messen versuchen, nicht automa-

tisch miteinander korreliert sein. Daher ist grundsätzlich zu hinterfragen, welche Vergleiche von Indikatoren einen Zusammenhang erwarten lassen.

Im Folgenden ordne ich Wissenschaftspreise in die Typologie von Forschungsleistungen und Indikatoren ein. Auf Basis dieser Kategorisierung kann abgeleitet werden, welche Indikatoren ähnliche Leistungsanstrengungen messen wie der Preis der Fritz Thyssen Stiftung und daher vermutlich miteinander korreliert sind. Forschungsleistungen lassen sich durch die Merkmale *Forschungshandeln*, *Nachwuchsförderung*, *Wissenschaftliche Infrastruktur* und *Wissenstransfer*² erfassen (Tabelle 1³). Ergebnisse des Forschungshandelns sind Publikationen und eingeworbene Drittmittel (Wissenschaftsrat 2008: 16ff.). Die zugehörigen Merkmalsausprägungen von Forschungshandeln und Nachwuchsförderung lassen sich wiederum in den Dimensionen *Originalität*, *Wirkung*, *Effektivität*, *Quantität* und *Effizienz* bewerten. Quantitative Indikatoren geben Auskunft über reine Häufigkeiten der Ausprägungen. Setzt man diese in Relation zu den eingesetzten Mitteln, erhält man einen Effizienzindikator. In den Dimensionen Effektivität und Wirkung wird die Forschungsleistung hinsichtlich definierter Zielgrößen (Kuhlmann, Heinze 2004: 56) und dem Beitrag zur Entwicklung der Wissenschaft (Wissenschaftsrat 2008: 19) evaluiert. Die Bewertungen in den drei erstgenannten Dimensionen sind Prozesse, die »in der Zeit veränderlich und in ihrer [...] aktuellen Gestalt unaufhebbar an das disziplinäre (kognitive) und professionelle (soziale) Leben einer wissenschaftlichen Gemeinschaft gebunden« sind (Hornbostel 1997: 187). Die Bewertungen von Originalität, Effektivität und Wirkung können daher zu einem bestimmten Zeitpunkt in einer bestimmten Profession höchstens intersubjektive Übereinstimmung finden und werden im Gegensatz zu den Dimensionen Quantität und Effizienz niemals objektiv zu bestimmen sein. *Qualitative*

2 Die Kategorisierung der Wissenschaftsindikatoren ist auf Basis der Ausführungen des Wissenschaftsrates (2008), Jansen et al. (2007) sowie Kuhlmann und Heinze (2004) vorgenommen worden und erhebt nicht den Anspruch auf Vollständigkeit. Eine ausführliche Erörterung zu Wissenschaftsindikatoren ist in Hornbostel (1997) sowie in Weingart und Winterhager (1984) zu finden. Die Validität und Reliabilität einzelner Indikatoren wird zum Teil ausgiebig diskutiert. Zu den meistverwendeten und zugleich umstrittensten Verfahren der Messung von Forschungsleistungen zählen bibliometrische Analyseverfahren.

3 Alle Tabellen und Abbildungen befinden sich im Anhang zu diesem Beitrag.

Indikatoren⁴ sind zum Beispiel (nach dem Publikationstyp und der internationalen Ausrichtung) gewichtete Publikationsanalysen, Zitationen⁵ und Wissenschaftspreise.⁶

In der Evaluationsforschung werden zusätzlich Engagements zur Aufrechterhaltung der wissenschaftlichen Infrastruktur und Wissens- und Technologietransfer als Merkmale von Forschungsleistungen identifiziert. Damit wird nicht nur das Forschungshandeln innerhalb des Wissenschaftssystems erfasst, sondern auch Anstrengungen zum Erhalt und Ausbau dieses Systems wie Herausgeberschaften von Fachzeitschriften und Aktivitäten, die auf einen Austausch von Informationen zwischen dem Wissenschaftssystem und anderen Systemen der Gesellschaft zielen, werden festgehalten.

Die vielen wissenschaftlichen Leistungsmerkmale können auf drei Arten miteinander verglichen werden. Entweder können unterschiedliche Hauptmerkmale von Forschungsleistungen wie Forschungshandeln und

4 Um die Darstellung verschiedener Wissenschaftsindikatoren knapp zu halten, verzichte ich auf eine nach den einzelnen qualitativen Dimensionen ausdifferenzierte Definition von Indikatoren. Interessierte Leserinnen und Leser finden hierzu in den Publikationen zum Forschungsrating der Soziologie des Wissenschaftsrates (2008) eine detaillierte Darstellung.

5 Meyer (2000) weist auf grundlegende Probleme von Zitationen als qualitative Publikationsindikatoren hin. Neben Zitierzirkeln, Selbstzitationen und Mehrfachpublikationen ähnlicher Inhalte ist eine große Schwäche in der mangelnden hermeneutischen Auslegung der Zitate zu sehen. Ablehnende Erwähnungen gehen ausnahmslos positiv in zitationsbasierte Indikatoren ein. In einer Untersuchung von Stigler und Friedland (1975) machten diese immerhin zehn Prozent der untersuchten Zitate aus. Münch (2007: 234) weist darauf hin, dass neben der tatsächlichen Qualität von Forschungsbeiträgen eine Vielzahl an anderen Faktoren für erste Rezeptionserfolge existiert. Der Anfangserfolg ist im erheblichen Maße eine soziale Konstruktion. Soziales Kapital in Form der Zusammenarbeit oder der Koautorenschaft mit etablierten Forschern ermöglicht es jungen Forschern, die ersten Hürden der Nichtbeachtung aufgrund des Angebotsüberschusses auf dem Zeitschriftenmarkt zu überwinden. Neben diesen methodischen Problemen existiert auch eine Reihe von technischen Problemen bei der richtigen Erfassung der Zitierhäufigkeiten (van Raan 2005: 136).

6 Wissenschaftspreise nehmen stark selektive Leistungsbewertungen im Bereich der Spitzenforschung vor und sind nur bedingt als Wissenschaftsindikatoren geeignet, da sie keine vollständige Reihung der Güte aller betrachteten Forschungsleistungen vornehmen (Hornbostel 1997: 208 ff.). Dennoch sind Wissenschaftspreise eine Form der Konstruktion von wissenschaftlicher Exzellenz (Alber et al. 2010). Daher wird der Preis der Fritz Thyssen Stiftung nachfolgend allgemein als qualitativer Indikator und nicht als qualitativer Wissenschaftsindikator bezeichnet.

Nachwuchsförderung, verschiedene Ausprägungen der Hauptmerkmale wie Drittmittelwerbung und Publikationen oder unterschiedliche Dimensionen einer Ausprägung wie qualitative und quantitative Publikationsindikatoren miteinander verglichen werden.

Der Vergleich der unterschiedlichen Hauptmerkmale erscheint auf Basis von Querschnittsdaten als schwierig, da aufgrund der Knappheit der Ressource Zeit Forschungshandeln (Publikationen und Drittmittel), Nachwuchsförderung, Engagement zur Erhaltung der wissenschaftlichen Infrastruktur und Wissenstransfer zu einem bestimmten Zeitpunkt in Konkurrenz zueinander stehen. Aus den gleichen Gründen sind Vergleiche einzelner Ausprägungen dieser Hauptmerkmale wie Publikationsoutput und Drittmittel ebenfalls problematisch. Naheliegender erscheinen Längsschnittanalysen, in denen zum Beispiel untersucht wird, ob ausgeprägte Publikationstätigkeiten der Vergangenheit spätere Tätigkeiten im Bereich der Erhaltung und des Ausbaus der wissenschaftlichen Infrastruktur oder die Drittmittelwerbung positiv beeinflussen.

Eine weitere Vergleichsmöglichkeit ergibt sich aus der Mehrdimensionalität des Wissenschaftsindikators Publikationen. Neben qualitativen Publikationsindikatoren wie Wissenschaftspreisen und dem Forschungsrating des Wissenschaftsrates existieren quantitative Publikationsindikatoren wie die Anzahl an Publikationen. Da eine ausgeprägte Publikationstätigkeit nicht notwendigerweise mit qualitativ hochwertiger Forschung einhergeht, können qualitative und quantitative Publikationsindikatoren unterschiedliche Evaluationsergebnisse generieren. Fruchtbarer erscheint die Analyse des Zusammenhangs mehrerer qualitativer Wissenschaftsindikatoren, da diese versuchen, das Gleiche zu messen. So hat der Wissenschaftsrat den Publikationsoutput von Forschungseinrichtungen im Bereich der Soziologie hinsichtlich seiner Originalität, Wirkung und Effektivität evaluiert. Die Anzahl publizierter Beiträge in hochrangigen Zeitschriften ist ebenfalls ein Indiz für die Güte von Forschungsleistungen.

Da der *Thyssen Preis* ein qualitativer Indikator im Bereich der Publikationen ist, vermute ich, dass dieser mit anderen qualitativen Publikationsindikatoren stark und mit quantitativen Publikationsindikatoren geringer beziehungsweise nicht korreliert ist. Von den alternativen qualitativen Publikationsindikatoren werden in der nachfolgenden empirischen Analyse Zitationen auf der Mikroebene der Aufsätze und auf der Aggregationsebene der Institutionen Leistungsindikatoren des Wissenschaftsrates aus dem Forschungsrating 2008 eingebunden:

H1: Prämierte Aufsätze werden häufiger zitiert als nominierte Aufsätze und nominierte Beiträge werden öfter zitiert als nicht nominierte Beiträge.

H2: Forschungseinrichtungen, die vom Wissenschaftsrat als forschungsstarke Institutionen eingeschätzt werden, haben größere Erfolge im Prämierungsverfahren aufzuweisen als forschungsschwache Institutionen.

Im Vergleich zu den qualitativen Indikatoren erfassen die vom CHE publizierten Publikationskennziffern aufgrund der verwendeten Erhebungs- und Gewichtungsverfahren eher die quantitative Dimension der Publikationskultur in der Forschung (Münch 2007: 266f.). Daher vermute ich, dass die Korrelation wesentlich geringer ausfällt als bei den qualitativen Indikatoren beziehungsweise dass sie gar nicht existiert:

H3: Forschungseinrichtungen, die Erfolge im Prämierungsverfahren aufweisen, werden vom Centrum für Hochschulentwicklung nicht notwendigerweise als publikationsstarke Institutionen identifiziert.

Methoden und Daten

Erfolgsgrößen des Prämierungsverfahrens und Rezeptionen

Der Zusammenhang von Zitationen und den Nominierungs- und Prämierungserfolgen wird auf der Mikroebene der Aufsätze untersucht. Von einer Zitationsanalyse auf der Aggregatebene der Autoren wie in Alber et al. (2010) nehme ich Abstand. Einerseits ist die Ermittlung der autorenspezifischen Zitationszahlen mit einem wesentlich höheren administrativen Aufwand verbunden. Andererseits wird der Zusammenhang zwischen Zitationen und dem Erfolg im Prämierungsverfahren auf der Mikroebene der Artikel unverzerrter erfasst, da durch jede Aggregation Informationen verloren gehen.

Der Primärdatensatz enthält alle 5.285 Zeitschriftenbeiträge, die am Prämierungsverfahren zwischen den Jahren 1981 und 2006 beteiligt waren, mit Angaben zur Affiliation der Autoren, zum Erscheinungsjahr der Aufsätze und zu den Zeitschriften, in denen die Aufsätze publiziert wurden. Darüber hinaus ist für jeden Aufsatz der Prämierungsstatus mit den Ausprägungen 1. Preis, 2. Preis, 3. Preis, nur nominiert (aber nicht prämiert) und nicht nominiert erfasst.

Als Indikator für die Sichtbarkeit eines Artikels werden die in Google Scholar gelisteten Zitationen genutzt. Google Scholar ist ein Internet basierter Suchdienst für wissenschaftliche Dokumente, der auf einen umfangreichen Datenpool von Verlagen, Organisationen und wissenschaftlichen Datenbanken zurückgreift.⁷ Die Leistungsfähigkeit von Google Scholar wird aktuell in der wissenschaftlichen Literatur kritisch diskutiert. Trotz einiger Schwächen wird Google Scholar insbesondere in der deutschsprachigen Szientometrie gegenüber dem Social Sciences Citation Index (SSCI) der Vorzug gegeben.⁸ Ein wesentlicher Nachteil von Google Scholar ist, dass im Gegensatz zum SSCI im Suchalgorithmus auch nicht-wissenschaftliche Literatur berücksichtigt wird. Dieser Nachteil wird aber dadurch aufgewogen, dass Google Scholar der diversifizierten Publikationskultur in der deutschsprachigen Sozialwissenschaft stärker Rechnung trägt. Während die Zitationen im SSCI nur auf Basis von Zeitschriftenbeiträgen ermittelt werden, erfasst Google Scholar auch Zitationen aus Büchern, Monographien und Sammelwerken. Ein weiterer Nachteil des SSCI ist die unzureichende Wahrnehmung deutschsprachiger Zeitschriften. Von den 17 am Prämierungsverfahren beteiligten Zeitschriften werden derzeit nur sieben Zeitschriften im SSCI ausgewertet. Auch wenn nicht jeder deutschsprachige Zeitschriftenbeitrag von Google Scholar erfasst wird, ist die Abdeckung aufgrund der Berücksichtigung aller hier relevanten Zeitschriften wesentlich höher.

Im Frühjahr 2010 wurden innerhalb von zwei Wochen für jeden Artikel die in Google-Scholar geführten Zitationen erhoben. Da die Artikel zu unterschiedlichen Zeitpunkten publiziert wurden, können die Zitationsmaße nur unter Berücksichtigung des Alters der Artikel miteinander verglichen werden. Eine einfache Lösung ist die Berechnung der durchschnittlichen Zitationen pro Jahr. Gegen diese Vorgehensweise sprechen jedoch die Ergebnisse der Szientometrie, der zufolge die Anzahl an Zitationen pro Artikel im Zeitablauf nicht konstant ist (Price 1976; Avramescu 1979). Ge-

⁷ Eine ausführlichere Beschreibung von Google Scholar findet sich in Paludkiewicz, Wohlrabe 2010.

⁸ Weiterführende Literatur zur Nutzung von Google Scholar in der wissenschaftlichen Evaluationsforschung findet sich unter anderem in Müller 2010; Breuer 2009; Harzing, van der Wal 2008; 2009 oder Meho, Yang 2007.

mäß dem Matthäus-Prinzip stärken gegenwärtige Zitationen die Sichtbarkeit eines Artikels und erhöhen zukünftige Zitationen. Dieser fortwährende Prozess führt zu einem überproportionalen Anstieg der Zitationszahlen über die Zeit. Diesem kumulativen Prozess wirkt jedoch die Obsoleszenz von Wissen entgegen, die ein exponentielles Wachstum der Zitationszahlen verhindert. Die Berücksichtigung beider Effekte lässt erwarten, dass die jährlichen Zitationen pro Artikel zunächst ansteigen und dann fallen.

Die empirische Forschung bestätigt im Allgemeinen den soeben beschriebenen Zitationszyklus. Demnach sind die Zitationszahlen kurz nach dem Erscheinen eines Aufsatzes gering, erreichen nach einem schnellen Anstieg einen Höhepunkt und nehmen in den Folgejahren dann langsam ab. Mehrere Studien weisen darauf hin, dass sich diese Verlaufsform zwischen zitationsschwachen und zitationsstarken Forschungsbeiträgen unterscheidet. Zitationsstarke Beiträge erreichen demnach den Höhepunkt später und altern danach langsamer als zitationsschwächere Beiträge.⁹

Da die Zitationen im Jahr 2010 erhoben wurden, ist die Modellierung der Entwicklung der Zitationszahlen über einen Zeitraum von 29 Jahren erforderlich. In Abbildung 1 sind drei unterschiedliche Zitationszyklen dargestellt. Jeder Typ zeigt für einen 29 Jahre alten Artikel die Verteilung der gesamten Anzahl an Zitationen auf die einzelnen Jahre, $c(t)$. Während Typ 1 zeitunabhängige Zitationsraten darstellt, beschreiben Typ 2 und 3 zeitabhängige Zitationszyklen. Gemäß den Ergebnissen der empirischen Forschung entspricht Typ 3 dem Zyklus zitationsstarker Beiträge. Die Zitationen erreichen relativ spät (nach 11 Jahren) ihren Höhepunkt und nehmen danach langsam ab. Zitationsschwache Beiträge (Typ 2) erreichen ihren Hochpunkt eher (nach sechs Jahren) und altern relativ schnell.¹⁰

Die Ermittlung der durchschnittlichen Zitationszahlen (Typ 1) hätte zur Folge, dass bei den hier modellierten Zitationszyklen die Anzahl der Zitationen von jüngeren Artikeln tendenziell überschätzt und die der älte-

⁹ Siehe zum Beispiel Walters 2011; Aksnes 2003; Cano, Lind 1991; Aversa 1985 oder Price 1976.

¹⁰ Bei der Modellierung der Verläufe stütze ich mich auf die Ergebnisse von Walters (2011). Walters untersucht die Zitationszyklen von zitationsstarken und -schwachen Zeitschriftenbeiträgen im Bereich der Psychologie zwischen den Jahren 1985 und 2010. Die weniger erfolgreichen Artikel weisen einen Hochpunkt nach ungefähr fünf und die erfolgreichen Artikel einen Hochpunkt nach 11 Jahren auf.

ren Artikel unterschätzt wird. In der weiteren Analyse wird daher für zitationsstarke Beiträge ein Zyklus des Typs 3 und für zitationsschwache Aufsätze ein Verlauf des Typs 2 unterstellt. Eine Publikation wird als zitationsstark eingestuft, wenn deren Zitationen höher als das 2,5-fache der durchschnittlichen Zitationen gleichaltriger Aufsätze sind.¹¹

Aggregierter Prämierungserfolg, CHE-Daten und Daten des Wissenschaftsrates

Auf der Institutionenebene wird der Prämierungs- und Nominierungserfolg der Forschungseinrichtungen mit den Ergebnissen des Forschungsratings des Wissenschaftsrates und den Hochschulrankings des Centrums für Hochschulentwicklung verglichen. Solche Vergleiche auf der Makroebene sind mit methodischen Problemen behaftet. So unterscheiden sich zum Beispiel die berücksichtigten Publikationsmedien im Thyssen-Sample und in den Hochschulrankings des CHE. Ebenfalls problematisch ist die Nutzung der Publikationskennziffern des CHE. Die vom CHE verwendeten Erhebungs- und Gewichtungsverfahren rufen regelmäßig Kritik aus der Fachgemeinschaft an der Güte der Publikationskennziffern hervor.¹² Darüber hinaus werden trotz der Dominanz soziologischer Beiträge auch Zeitschriftenaufsätze aus der Politikwissenschaft, Geschichte und Psychologie im Prämierungsverfahren berücksichtigt. Die Hochschulrankings des CHE und das Forschungsrating des Wissenschaftsrates beziehen sich hingegen nur auf den Fachbereich Soziologie. Die hier abgeleiteten Ergebnisse sollten daher nur als Annäherung an tatsächliche Gesetzmäßigkeiten verstanden werden.

11 Diese Vorgehensweise ist den Arbeiten von Aksens (2003) und Glänzel et al. (1995) entlehnt. Die unumgänglichen arbiträren Annahmen über den Verlauf der Zitationszyklen und über die Einteilung der Artikel in zitationsschwache und -starke Beiträge sind meines Erachtens unproblematisch, da die empirischen Untersuchungsergebnisse relativ robust gegenüber diesen Annahmen sind.

12 Vgl. Ausführungen der vorangegangenen theoretischen Diskussion, Münch (2007: 264 ff.) zur Verwendung der CHE-Daten oder Gläser (2006) sowie Liebeskind und Ludwig-Mayerhofer (2005) zu allgemeinen Problemen bibliometrischer Verfahren zur Messung von Forschungsleistungen.

Vom CHE werden die Publikationszahlen verwendet, die in den drei Hochschulrankings der Jahre 2002, 2005 und 2008 für den Fachbereich Soziologie veröffentlicht worden sind (Berghoff et al. 2002; 2005; 2008). Neben diesen Absolutwerten enthalten die Rankings mit den Publikationen pro Professor/Professorin (Ranking 2002, 2005) beziehungsweise Publikationen pro Wissenschaftler/Wissenschaftlerin (Ranking 2008) ebenfalls relative Erfolgsmaße. Darüber hinaus ist zu beachten, dass es sich in den Rankings der Jahre 2002 und 2005 um nach Art und Umfang gewichtete Publikationen und im Ranking 2008 um die ungewichtete Anzahl aller internationalen Publikationen handelt.

Aus dem Forschungsrating Soziologie des Wissenschaftsrates (2008) werden die Evaluationsergebnisse des Forschungshandelns in den Dimensionen *Originalität*, *Effektivität* und *Effizienz* genutzt. Unter Ersterem soll die Originalität und die wissenschaftliche Bedeutung erbrachter Forschungsleistungen erfasst werden. Effektivität ist ein Indikator für den Beitrag einer Forschungseinrichtung zur Entwicklung der Wissenschaft und Effizienz beschreibt die Effektivität einer Einrichtung in Relation zum eingesetzten wissenschaftlichen Personal. Aufgrund der methodischen Vorgehensweise interpretiert der Wissenschaftsrat die erfasste Originalität und die Effizienz als relative Erfolgsgrößen und die Effektivität als absolute Erfolgsgröße.

Die drei Leistungsindikatoren des Wissenschaftsrates wurden auf einer Skala von *nicht befriedigend* bis *exzellent* erfasst. Für die vorliegende Untersuchung habe ich diese Angaben in eine metrische Skalierung von 1 für *nicht befriedigend* bis 5 für *exzellent* transformiert.

Der Erfolg einer Einrichtung im Prämierungsverfahren der Fritz Thyssen Stiftung wird analog zu Alber et al. (2010) anhand der *Thyssen-Punkte* durch die gewichtete Summe der erhaltenen Preise und Nominierungen gemessen. Erste Preise erhalten fünf Punkte, zweite Preise vier Punkte, dritte Preise drei Punkte und nur Nominierungen (ohne Prämierung) einen Punkt. Dividiert man die Thyssen-Punkte einer Institution durch die Anzahl der publizierten Artikel der Institution, erhält man ein relatives Erfolgsmaß.

Um die Vergleichbarkeit des Nominierungs- und Prämierungserfolgs mit den Kennziffern des CHE und des Wissenschaftsrates zu erhöhen, wurden die *Thyssen-Punkte* für die jeweiligen Zeitperioden der einzelnen Hochschulrankings beziehungsweise des Forschungsratings berechnet. Für das CHE-Ranking 2002 sind die Jahre 1998 bis 2000, für das CHE-Ranking 2005 die Jahre 2001 bis 2003, für das Ranking 2008 die Jahre 2004 bis 2006 und für das WR-Rating die Jahre 2001 bis 2005 relevant. Die Anzahl

der berücksichtigten Forschungseinrichtungen richtet sich nach den Grundgesamtheiten des Wissenschaftsrates und des CHE. Im CHE-Ranking 2002 wurden 45, im CHE-Ranking 2005 54, im CHE-Ranking 2008 55 Hochschulen und im WR-Rating wurden 51 Hochschulen und drei außeruniversitäre Forschungseinrichtungen evaluiert.

Datenanalyse

Erfolgsgrößen des Prämierungsverfahrens und Rezeptionen

In Tabelle 2 sind für die 1980er, 1990er und 2000er Jahre einige Maßzahlen der Verteilungen der adjustierten Zitationshäufigkeiten pro Jahr abgebildet. Die Zitationsraten sind mit Durchschnittswerten zwischen 0,48 in den 1980er, 0,79 in den 1990er und 1,24 in den 2000er Jahren relativ niedrig. Die Abnahme im Zeitverlauf verdeutlicht dabei, dass der Durchschnittsartikel mit der Zeit an Aktualität und an Relevanz verliert. Standardabweichungen, Minima und Maxima der einzelnen Perioden weisen auf eine starke Varianz der Zitationen hin. Die sehr hohen Krümmungswerte zwischen 3,76 und 11,30 deuten auf stark rechtsschief verteilte Zitationen hin.

Anhand der Box-Plots der Abbildung 2 ist deutlich zu erkennen, dass sich die nach dem Prämierungsstatus differenzierten Verteilungen der adjustierten Zitationshäufigkeiten der Artikel deutlich unterscheiden. Charakteristisch für alle Verteilungen ist die starke Konzentration auf niedrige Zitationshäufigkeiten. Dabei liegen in der Gruppe der nicht nominierten Artikel 50 Prozent aller Artikel um den Median zwischen 0 und 0,86 und in der Gruppe der ersten Preise zwischen 0,87 und 4,77. Die Gruppen der nur nominierten und nicht nominierten Artikel weisen viele Ausreißer auf und enthalten die Artikel, die relativ am häufigsten zitiert wurden. Trotzdem ist mit Ausnahme der zweiten Preise der Median in höheren Gruppen größer.

Die in der Tabelle 3 abgebildeten Mittelwerte entsprechen meiner Vermutung, dass nicht nominierte Aufsätze im Vergleich zu allen anderen Gruppen und nur nominierte Aufsätze im Vergleich zu den prämierten Aufsätzen seltener zitiert werden. Obwohl die absoluten Unterschiede gering ausfallen, sind die relativen Mittelwertdifferenzen erheblich. So wurden zum Beispiel Aufsätze, die mit einem 3. Preis prämiert wurden, durchschnittlich doppelt so häufig zitiert wie nicht nominierte Aufsätze. Beiträge, die mit einem 1. Preis prämiert wurden, haben viermal höhere Zitatio-

nen als nicht nominierte Aufsätze. Besonders bemerkenswert ist meines Erachtens, dass die Mittelwerte sogar in den einzelnen Preisstufen zunehmen. Somit spiegelt sich die Ordinalität der Preisstufen in den Zitationshäufigkeiten wider.

Zur Überprüfung der Signifikanz der Mittelwertunterschiede der einzelnen Gruppen habe ich eine Varianzanalyse mit einem Kruskal-Wallis-Rangsummen-Test durchgeführt (Tabelle 4).¹³ Insgesamt bestätigt die Varianzanalyse, dass die Mittelwertunterschiede in den meisten Fällen hoch signifikant sind.¹⁴ Wie ich eingangs vermutet habe (H1), werden nicht nominierte Aufsätze im Vergleich zu nur nominierten Aufsätzen und nur nominierte Aufsätze im Vergleich zu prämierten Aufsätzen signifikant weniger zitiert. Richtet man den Fokus auf die Zitationshäufigkeiten der Gruppen der einzelnen Preisstufen, ist zu erkennen, dass die Unterschiede in den Zitationsraten von Aufsätzen, die mit einem dritten und zweiten Preis prämiert worden sind, nicht signifikant sind. Dies ist vor allem auf die hohe Varianz der Zitationsraten von Aufsätzen zurückzuführen, die mit einem zweiten Preis prämiert worden sind. Aufsätze, die mit einem ersten Preis prämiert worden sind, werden im Gegensatz dazu signifikant häufiger zitiert als Aufsätze, die einen zweiten oder dritten Preis erhalten haben.

Alber und Fliegner (2010) zeigen, dass die Kölner Zeitschrift für Soziologie und Sozialpsychologie, die Zeitschrift für Soziologie, das Berliner Journal für Soziologie, die Politische Vierteljahresschrift, Leviathan und die Schweizer Zeitschrift für Politikwissenschaft die größten absoluten und relativen Prämierungserfolge aufweisen. Da diese Zeitschriften zu den Kernzeitschriften der deutschsprachigen Sozialwissenschaft zählen, genießen ihre Aufsätze eine große Sichtbarkeit. Dies zeigen auch die durchschnittlichen adjustierten Zitationen der Aufsätze, die in der hier genannten Erfolgsgruppe 1.20 und in der Gruppe der übrigen Zeitschriften 0.47 betragen. Da nicht nominierte Beiträge überproportional häufig in den zi-

13 Eine Varianzanalyse mit Bonferroni Anpassung ist nicht möglich, da ein durchgeführter Levene-Test die Annahme der Varianzhomogenität zwischen den Gruppen verwirft. Angesichts der in Tabelle 3 abgebildeten Standardabweichungen ist dieses Ergebnis nicht überraschend.

14 Dieses Ergebnis ist robust gegenüber dem angenommenen Zitationszyklus. Analysen mit alternativen Zitationszyklen weisen nur minimalste Abweichungen in den Mittelwertdifferenzen auf.

tationsschwachen Zeitschriften veröffentlicht sind, werden möglicherweise die Mittelwertdifferenzen zwischen den nicht nominierten und den prämierten Aufsätzen überschätzt. Dies bestätigt eine Regression mit robust geschätzten Standardfehlern, in der auf die adjustierten Zitationen die Dummies für den jeweiligen Prämierungsstatus (Referenzgruppe: nicht nominierte Aufsätze) und ein Dummy für die hier definierten Erfolgszeitschriften regressiert werden:¹⁵

$$\begin{aligned}
 (\text{adj.}) \text{ Zitationen} &= 0,441 + 0,382 \times \text{nur nom.} + 0,644 \times 3. \text{ Preis} + 1,472 \times 2. \text{ Preis} \\
 &\quad (0,000) \quad (0,000) \qquad\qquad (0,071) \qquad\qquad (0,069) \\
 &+ 2,128 \times 1. \text{ Preis} + 0,719 \times \text{Erfolgszeitschrift} \\
 &\quad (0,000) \qquad\qquad (0,000)
 \end{aligned}$$

Der Vergleich der Regressionskoeffizienten der Preisstufen mit den in Tabelle 4 (Spalte »nicht nom.«) ausgewiesenen Ergebnissen zeigt, dass bei Kontrolle des Publikationsmediums die Mittelwertdifferenzen etwas geringer ausfallen. So weisen zum Beispiel Aufsätze, die mit dem ersten Preis prämiert worden sind, durchschnittlich 2,128 anstatt 2,365 Zitationen mehr aus als nicht prämierte Aufsätze. Insgesamt sind die Abweichungen jedoch relativ gering, so dass die zuvor festgestellten Ergebnisse bestätigt werden: mit zunehmendem Prämierungsstatus steigen die Zitationen.

Umgeklärt ist die Frage, ob es sich bei diesem Zusammenhang um eine reine Korrelation handelt oder ob ein kausaler Effekt existiert. Da zwischen der Veröffentlichung eines Aufsatzes und der potentiellen Nominierung und Prämierung etwa sechs bis 18 Monate liegen, kann ein kausaler Effekt der Zitationen auf die Erfolgsgrößen des Prämierungsverfahrens ausgeschlossen werden. Dieses Zeitfenster ist zu klein, als dass ein zu prämierender oder zu nominierender Aufsatz von einer größeren Anzahl an anderen Publikationen zitiert wird und diese zitierenden Aufsätze von der Jury und den Zeitschriftenherausgebern wahrgenommen werden.

Ein umgekehrter kausaler Zusammenhang kann jedoch auf der Prämierungsebene (prämiert versus nicht prämiert) nicht ausgeschlossen werden. Zum einen stärkt eine Prämierung die Sichtbarkeit eines Aufsatzes und zum anderen stuft die Jury der Fritz Thyssen Stiftung den Beitrag als he-

¹⁵ Abgebildet sind nicht standardisierte Koeffizienten und p-Werte in den Klammern.

rausragende Forschungsleistung ein. Da die Veröffentlichung eines Aufsatzes und das Prämierungsverfahren zeitlich zu nah beieinander liegen, wird eine abschließende empirische Kausalanalyse niemals möglich sein. Auf der Nominierungsebene (nicht nominiert versus nur nominiert) kann ein kausaler Zusammenhang ausgeschlossen werden. In den jährlichen Mitteilung zur Preisvergabe der Fritz Thyssen Stiftung werden einzig die prämierten Aufsätze genannt und daher sind der breiten Wissenschaftsgemeinde die nominierten Aufsätze nicht bekannt.

Preis der Fritz Thyssen Stiftung und qualitative Leistungsindikatoren des Wissenschaftsrates

Der Wissenschaftsrat hat in seinem Forschungsrating des Fachs Soziologie mit den Indikatoren *Originalität*, *Effizienz* und *Effektivität* drei weitere qualitative Leistungsindikatoren des Forschungshandelns erhoben. Die empirische Untersuchung bestätigt unsere Vermutung des positiven Zusammenhangs zwischen diesen Leistungsmerkmalen und den Erfolgsgrößen des Prämierungsverfahrens (H2). Dabei ist die Steigung der Regressionsgeraden bei den absoluten Werten der Effektivität und der gesamten Thyssen-Punkte relativ steil und die Korrelation beträgt $r=0,461$ (Abbildung 3a). Bei den relativen Werten verlaufen die Regressionsgeraden etwas flacher. Die Korrelation zwischen Originalität und Thyssen-Punkten pro Publikation beträgt $r=0,275$ und zwischen Effizienz und Thyssen-Punkten pro Publikation $r=0,251$ (3b, 3c). Trotz der relativ schwachen Korrelationen sind die Zusammenhänge hoch signifikant. Forschungseinrichtungen, deren Forschungsleistungen nach Meinung des Wissenschaftsrates tendenziell einen hohen Grad an Originalität aufweisen und deren Forschungsanstrengungen einen hohen Beitrag zur Entwicklung der Wissenschaft leisten, zeichnen sich demnach ebenfalls durch leicht höhere relative Erfolge im Prämierungsverfahren der Fritz Thyssen Stiftung aus.

Ein positiver Zusammenhang zwischen den absoluten Leistungsindikatoren überrascht nicht. Aufgrund der methodischen Ermittlung des Indikators *Effektivität* schneiden größere Forschungseinrichtungen in dieser Kategorie tendenziell besser ab als kleinere Einrichtungen (Wissenschaftsrat 2008: 10). Gleichzeitig haben größere Einrichtungen in der Regel höhere absolute Publikationszahlen und dadurch höhere absolute Erfolgchancen im Prämierungsverfahren der Fritz Thyssen Stiftung.

Die in der Literatur beschriebene mehrfache *zwei-Welten-Lehre* ist aus dem Befund abgeleitet, dass mehrere *relative* Leistungsindikatoren auseinanderklaffen. Dieses Phänomen ist in der vorliegenden Analyse qualitativer Leistungsindikatoren nicht zu beobachten. Die hoch signifikante Korrelation der relativen Erfolgsindikatoren stützt die theoretischen Überlegungen, dass sowohl der Wissenschaftspreis der Fritz Thyssen Stiftung als auch die Leistungsindikatoren des Wissenschaftsrates *Originalität* und *Effizienz* die qualitative Dimension von Forschungshandeln erfassen und daher ein Zusammenhang zwischen diesen Leistungsmerkmalen besteht.

Da der Wissenschaftsrat Forschungspreise bei der Ermittlung der Effizienz und Effektivität explizit berücksichtigt, ist der Preis der Fritz Thyssen Stiftung ein bestimmendes Element dieser Indikatoren. Reputierungen in beiden Dimensionen sind daher nicht als unabhängige Erfolge zu werten. Im Gegensatz dazu werden Wissenschaftspreise zur Ermittlung der Originalität nicht herangezogen. Forschungseinrichtungen reputieren sich daher unabhängig in gleichem Maße durch den Wissenschaftspreis der Fritz Thyssen Stiftung und durch die Qualität von Forschungsleistungen.

Preis der Fritz Thyssen Stiftung und Publikationen

Die bivariate Analyse des Forschungsausgangs in Form von Publikationen und dem Erfolg im Prämierungsverfahren zeigt einen signifikant positiven Zusammenhang bei den Absolutwerten (Abbildungen 4a, 4c, 4e) und einen nur schwachen oder keinen Zusammenhang bei den relativen Werten (Abbildungen 4b, 4d, 4f). Während die Regressionsgeraden der Absolutwerte in allen Perioden eine moderate Steigung aufweisen und die Korrelation um 0.3 liegt, ist die Korrelation zwischen Thyssen-Punkten pro Publikation und Publikationen pro Professor/Professorin (Wissenschaftler/Wissenschaftlerin) mit $r=0,12$ zwischen den Jahren 2001 und 2003 sowie $r=0,09$ zwischen den Jahren 2004 und 2006 gering und zwischen den Jahren 1998 und 2000 sogar negativ ($r=-0,08$). Die Streudiagramme der relativen Werte zeigen in allen Perioden unsystematische Konstellationen aus Thyssen-Punkten und Publikationen. Es gibt Institute mit ausgeprägten Prämierungs- und Nominierungserfolgen und niedrigen Publikationszahlen und es gibt umgekehrt Einrichtungen, die viele Publikationen und nur wenige Nominierungen und Prämierungen zu verzeichnen haben.

Diese Ergebnisse stehen im Einklang mit Münchs (2007) Untersuchung, der in den Geistes- und Sozialwissenschaftlichen positive Korrelationen zwischen den gesamten verfügbaren Drittmitteln und allen Publikationen einer Forschungseinrichtung und tendenziell keinen Zusammenhang zwischen Drittmitteln pro Wissenschaftler und Publikationen pro Professor bzw. Professorin findet.

Der positive Zusammenhang zwischen den absoluten Leistungsindikatoren wird auf Größeneffekte zurückzuführen sein. Durch Drittmittel werden Mitarbeiter finanziert und mit einem Anstieg der Mitarbeiterzahlen einer Forschungseinrichtung geht in der Regel eine höhere absolute Produktivität einher (Münch 2007: 269), die die Publikationen im Sample des CHE und im Thyssen-Sample erhöht. Mit der Anzahl der Beiträge im Thyssen-Sample steigen die Chancen auf Nominierung und Prämierungen. Münch interpretiert den Zusammenhang zwischen den absoluten Leistungsindikatoren durch den von Merton (1957; 1968) beschriebenen, sich selbst verstärkenden Prozess der Akkumulation von Vorteilen. Forschungseinrichtungen können demnach einen Machtvorsprung bei der Einwerbung von Drittmitteln in eine Erhöhung der absoluten Publikationszahlen umsetzen. Die vorliegende Untersuchung zeigt nun, dass der Akkumulationsprozess dadurch verschärft wird, dass ein hoher absoluter Publikationsoutput in einen Vorsprung bei dem absoluten Erfolg im Prämierungsverfahren der Fritz Thyssen Stiftung umgewandelt werden kann.

Bei Kontrolle des Größeneffekts zeigt sich in Münchs und in der vorliegenden Untersuchung ein Auseinanderklaffen der relativen Leistungsindikatoren. In Bezug auf Drittmittel und Publikationen kann dies durch die Knappheit der Ressource Zeit oder durch das System der Drittmittelforschung erklärt werden, welches innovative Forschung nicht fördert und tendenziell publikationsunwürdige Forschungsergebnisse hervorbringt (Münch 2007: 277).

Wie vermutet, existiert auch zwischen Thyssen-Punkten pro Publikation und Publikationen pro Professor/in (Wissenschaftler/in) kein robuster Zusammenhang (H3). Interpretiert man dieses Ergebnis vor dem Hintergrund der aus der vorangegangenen theoretischen Diskussion bestimmten Heterogenität von Leistungsindikatoren, bekräftigt der fehlende Zusammenhang die Vermutung, dass Forschungspreise und Publikationen unterschiedliche Dimensionen des Forschungsoutputs messen, die nicht zwangsläufig miteinander einhergehen. Publikationen sind Indikatoren für die quantitative Dimension von Forschungshandeln, wohingegen der Preis der Fritz Thyssen Stiftung ein Indikator für die qualitative Dimension des Forschungshandelns ist.

Zusammenfassung

Reputation in der Forschung entsteht aus den unterschiedlichen Leistungen des wissenschaftlichen Handelns wie der Ausbildung des wissenschaftlichen Nachwuchses, der Pflege der wissenschaftlichen Infrastruktur, der Einwerbung von Drittmitteln, der Erschließung neuen Wissens und der Aufbereitung dieser neu gewonnenen Erkenntnisse in Form von Publikationen. Aufgrund der Heterogenität dieser Leistungsmerkmale zeichnen sich Forschungseinrichtungen und Wissenschaftler nicht notwendigerweise gleichzeitig in all diesen Reputationsdimensionen aus. Insbesondere in der Soziologie hat die Evaluationsforschung das Auseinanderklaffen von unterschiedlichen Leistungsindikatoren mehrfach festgestellt.

Weitgehend ungeklärt ist bisher die Frage, ob Wissenschaftspreise eine eigene Reputationsdimension bilden oder in engem Zusammenhang mit anderen wissenschaftlichen Leistungsmerkmalen stehen. Die vorliegende Untersuchung zeigt exemplarisch am Beispiel des Preises der Fritz Thyssen Stiftung, dass dieses Prämierungsverfahren keine völlig neue Dimension der Exzellenzzuschreibung bildet. Obwohl auf institutioneller Ebene kein Zusammenhang zwischen Publikationsoutput und den Erfolgsgrößen des Prämierungsverfahrens festzustellen ist, existiert ein Zusammenhang mit zwei weiteren Leistungsindikatoren. Zum einen werden prämierte Aufsätze häufiger zitiert als nur nominierte Aufsätze, und diese werden häufiger zitiert als nicht nominierte Aufsätze. Zum anderen existiert ein Zusammenhang zwischen der vom Wissenschaftsrat erhobenen Originalität, Effizienz und Effektivität von Forschungsleistungen der Forschungseinrichtungen und deren Erfolg im Prämierungsverfahren.

Die empirischen Ergebnisse bestätigen die Vermutung, dass der Zusammenhang zwischen zwei wissenschaftlichen Leistungsindikatoren stark davon abhängt, ob beide Faktoren die gleiche Dimension von Forschungsleistungen messen. Das Auseinanderklaffen von Prämierungserfolg und Publikationszahlen ist darauf zurückzuführen, dass die Anzahl an Publikationen die quantitative Dimension und der Preis der Fritz Thyssen Stiftung die qualitative Dimension der Publikationsaktivitäten misst. Der positive Zusammenhang zwischen den Leistungsmerkmalen des Wissenschaftsrates beziehungsweise Zitationen und dem Prämierungserfolg wird hingegen dadurch erklärt, dass die drei Merkmale als qualitative Indikatoren unter die gleiche Reputationsdimension subsumiert werden können.

Fraglich bleibt, ob die Korrelationen der hier betrachteten qualitativen Leistungsindikatoren unabhängig voneinander sind oder ob sie einander bedingen. Es ist nicht auszuschließen, dass der Preis der Fritz Thyssen Stiftung einen positiven Effekt auf die Zitationen und – aufgrund der methodischen Vorgehensweise des Wissenschaftsrates – auf die Wissenschaftsindikatoren Effizienz und Effektivität hat. In Unkenntnis des genauen Wirkungsmechanismus kann aber festgehalten werden, dass das Prämierungsverfahren der Fritz Thyssen Stiftung keine gänzlich neue Welt der wissenschaftlichen Reputation darstellt. Aufgrund der zum Teil nur schwachen Korrelationen scheint es für Evaluierungen im Wissenschaftsbetrieb aber unabdingbar zu sein, alle verfügbaren Indikatoren zu betrachten und dadurch die stark diversifizierte Welt wissenschaftlicher Leistungen in ihrer Gänze zu erfassen.

Literatur

- Aksens, D.W. 2003: Characteristics of Highly Cited Papers. *Research Evaluation*, 12. Jg., Heft 3, 159–170.
- Alber, J., Fliegner, F. 2010: Rezeption, Themen und Publikationsorgane der mit dem Preis der Fritz Thyssen Stiftung ausgezeichneten sozialwissenschaftlichen Aufsätze 1981–2006. *Soziologie*, 39. Jg., Heft 3, 286–309.
- Alber, J., Fliegner, F., Nerlich, T. 2010: Was bestimmt Exzellenz in den Sozialwissenschaften? Eine Analyse am Beispiel des Preises der Fritz Thyssen Stiftung für sozialwissenschaftliche Aufsätze 1981–2006. *Soziologie*, 39. Jg., Heft 2, 152–178.
- Allmendinger, J. 2002: Eine drei-Welten-Lehre wissenschaftlicher Reputation und ihre Messung. *Soziologie*, 31. Jg., Heft 3, 56–58.
- Aversa, E. 1985: Citation Patterns of Highly Cited Papers and their Relationship to Literature Aging: A Study of the Working Literature. *Scientometrics*, 7. Jg., Heft 3–6, 383–389.
- Avramescu, A. 1979: Actuality and Obsolescence of Scientific Literature. *Journal of the American Society for Information Science*, 30. Jg., Heft 5, 296–303.
- Berghoff, S., Federkeil, G., Giebisch, P., Hachmeister, C.-D., Hennings, M., Roessler, I., Ziegele, F. 2008: Das CHE-Forschungsranking deutscher Universitäten 2008 – *Soziologie* (2008). Arbeitspapier Nr. 114.
- Berghoff, S., Federkeil, G., Giebisch, P., Hachmeister, C.-D., Müller-Böling, D. 2002: Das Forschungsranking deutscher Universitäten – Analysen und Daten im Detail. Arbeitspapier Nr. 40.

- Berghoff, S., Federkeil, G., Giebisch, P., Hachmeister, C.-D., Müller-Böling, D. 2005: Das CHE-Forschungsranking deutscher Universitäten 2005. Arbeitspapier Nr. 70.
- Breuer, W. 2009: Google Scholar as a Means for Quantitative Evaluation of German Research Output in Business Administration: Some Preliminary Results. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1280033, Zugriff: 10/2011.
- Cano, V., Lind, N. 1991: Citation Life Cycles of Ten Citation Classics. *Scientometrics*, 22. Jg., Heft 2, 297–312.
- Endrweit, G. 2002: Wie misst man Reputation? Messtheoretische Überlegungen zu Jürgen Gerhards »Reputation in der deutschen Soziologie«. *Soziologie*, 31. Jg., Heft 4, 33–41.
- Gerhards, J. 2002: Reputation in der deutschen Soziologie – Zwei Welten. *Soziologie*, 31. Jg., Heft 2, 19–33.
- Glänzel, W., Rinia, E. J., Brocken, M. G. 1995: A Bibliometric Study on Highly-cited European Physics Papers in the 80s. *Research Evaluation*, 5. Jg., Heft 2, 113–122.
- Gläser, J., 2006. Fallstricke der Bibliometrie. *Soziologie*, 35. Jg., Heft 1, 42–51.
- Harzing, A.-W., van der Wal, R. 2008: Google Scholar as a new Source for Citation Analysis. *Ethics in Science and Environmental Politics*, 8. Jg., Heft 1, 61–73.
- Harzing, A.-W., van der Wal, R. 2009: A Google Scholar h-index for Journals: An Alternative Metric to Measure Journal Impact in Economics and Business. *Journal of the American Society for Information Science and Technology*, 60. Jg., Heft 1, 41–46.
- Hirschauer, S. 2004: Peer Review Verfahren auf dem Prüfstand. *Zeitschrift für Soziologie*, 33. Jg., Heft 1, 62–83.
- Hornbostel, S. 1997: *Wissenschaftsindikatoren. Bewertungen in der Wissenschaft.* Opladen: Westdeutscher Verlag.
- Jansen, D., Wals, A., Franke, K., Schmoch, U., Schubert, T. 2007: Drittmittel als Performanzindikator der Wissenschaftlichen Forschung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 59. Jg., Heft 1, 125–149.
- Kuhlmann, S., Heinze, T. 2004: Evaluation von Forschungsleistungen in Deutschland: Erzeuger und Bedarf. Teil I: Konzeptionelle Grundlagen. *Wissenschaftsrecht*, 27. Jg., 53–69.
- Liebeskind, U., Ludwig-Mayerhofer, W. 2005: Auf der Suche der Wunsch-Universität – im Stich gelassen. Anspruch und Wirklichkeit von Hochschulrankings. *Soziologie*, 34. Jg., Heft 4, 442–462.
- Meho, L. I., Yang, K. 2007: Impact of Data Sources on Citation Counts and Rankings of LIS Faculty: Web of Science, Scopus, and Google Scholar. *Journal of American Society for Information Science and Technology*, 58. Jg., Heft 13, 2105–2125.
- Merton, R. K. 1957: Science and Democratic Social Structure. In R. K. Merton, *Social Theory and Social Structure*, 3. Aufl., New York: Free Press of Glencoe.

- Merton, R. K. 1968: The Matthew Effect in Science. *Science*, 159. Jg., Heft 3810, 56–63.
- Meyer, D. 2000: Über die Arbeit wissenschaftlicher Zeitschriften in der Ökonomie. *Leviathan*, 28. Jg., Heft 1, 87–108.
- Müller, H. 2010: Wie valide ist das Handelsblatt-BWL-Ranking? Zeitschriften- und zitationsbasierte Personenrankings im Vergleich. *Betriebswirtschaftliche Forschung und Praxis*, 62. Jg., Heft 2, 150–164.
- Münch, R. 2007: Die akademische Elite. Frankfurt am Main: Suhrkamp.
- Paludkiewicz, K., Wohlrabe, K. 2010: Qualitätsanalyse von Zeitschriften in den Wirtschaftswissenschaften: Über Zitationsdatenbanken und Impaktfaktoren im Online-Zeitalter. *Ifo Schnelldienst*, 63. Jg., Heft 21, 18–28.
- Price, D. D. S. 1976: A General Theory of Bibliometric and other Cumulative Advantage Processes. *Journal of the American Society for Information Science*, 27. Jg., Heft 5–6, 292–306.
- Rammert, W. 2002: Die halbierte Reputation – eine grob fahrlässige und unfaire Rechnung! *Soziologie*, 31. Jg., Heft 3, 53–55.
- Schulz-Schaeffer, I. 2002: Publikationen zählen – Empirische Anmerkungen zum Publikations-Ranking und zur Reputationswelten-Lehre. *Soziologie*, 31. Jg., Heft 4, 42–55.
- Stigler, G., Friedland, C. 1975: The Citation Practices of Doctorates in Economics. *Journal of Political Economy*, 83. Jg., Heft 3, 477–507.
- Ursprung, H. 2003: Schneewittchen im Land der Klapperschlangen: Evaluation eines Evaluators. *Perspektiven der Wirtschaftspolitik*, 4. Jg., Heft 2, 177–189.
- van Raan, A. 2005: Fatal Attraction: Conceptual and Methodological Problems in the Ranking of Universities by Bibliometric Methods. *Scientometrics*, 62. Jg., Heft 1, 133–143.
- Walters, G. D. 2011: The Citation Life Cycle of Articles Published in 13 American Psychological Association Journals: A 25-year Longitudinal Analysis. *Journal of the American Society for Information Science and Technology*, 62. Jg., Heft 8, 1629–1636.
- Weingart, P., Winterhager, M. 1984: Die Vermessung der Forschung. Frankfurt am Main: Campus.
- Wissenschaftsrat 2008: Forschungsleistungen deutscher Universitäten und außer-universitärer Einrichtungen in der Soziologie. http://www.wissenschaftsrat.de/download/Forschungsrating/Dokumente/Pilotstude_Forschungsrating_Soziologie/pilot_ergeg_soziologie.pdf, Zugriff: 10/2011.

Anhang 1: Tabellen

Tabelle 1: Forschungsleistung und Indikatoren der sozialwissenschaftlichen Forschung (Auswahl)

Merkmal	Ausprägungen	Dimension	
		Originalität u. Wirkung/ Effektivität	Quantität
Forschungs- handeln	Publikationen	Publikationstyp (Zeitschriftenbeitrag mit/ohne Peer Review, Monographie, national/ international, graue Literatur) Sichtbarkeit (Zitationen, Rezensionen, Repliken) Expertenurteil (Wissenschaftspreise, Forschungsrating des WR)	Umfang der Publikationen, Anzahl an Promotionen/ Habilitationen und Volumen der eingeworbenen Drittmittel
Nachwuchs- förderung	Drittmittel Promotionen, Habilitationen	Auftraggeber (DFG, Stiftungen, Bundes-)Ministerien) Expertenurteil Benotung, Dauer, Rufe an Nachwuchswissenschaftler auf Stellen von Professuren	
Wissenschaftliche Infrastruktur	Herausgeberschaften in Fachzeitschriften, Ämter in anderen Einrichtungen, Tätigkeit in wissenschaftlichen Beratungsgremien		
Wissens- und Technologie- transfer	Ämter außerhalb der Wissenschaft, Dienstleistungen, Auftragsforschung und Gutachten, Kooperationen mit Unternehmen, Weiterbildungsangebote, Publikationen außerhalb des Wissenschaftsbetriebs wie zum Beispiel in Tageszeitungen oder Mitteilungen von kollektiven Akteuren wie Gewerkschaften und Arbeitgeberverbänden		Absolute Höhe / Volumen der Merkmalsausprägung in Relation der eingesetzten Mittel

Tabelle 2: Verteilung der adjustierten Zitationen pro Jahr

Period	N	Mean	Std.Dev.	Skewness	Min.	Max.
2000er	1.654	1,247	2,090	3,755	0	25,128
1990er	2.316	0,786	1,543	4,983	0	17,912
1980er	1.315	0,476	1,328	11,303	0	28,314
Gesamt	5.285	0,853	1,713	5,220	0	28,314

Tabelle 3: Verteilung der (adj.) Zitationen pro Jahr nach Prämierungsstatus

Präm.status	Mean	Std. Dev.	N
nicht nom.	0,795	1,576	4.633
nur nom.	1,112	2,208	559
3. Preis	1,540	2,339	40
2. Preis	2,371	4,511	31
1. Preis	3,160	2,962	22
Gesamt	0,853	1,713	5.285

Tabelle 4: Varianzanalyse der (adj.) Zitationsraten nach Prämierungsstaus

Präm.status	nicht nom.	nur nom.	3. Preis	2. Preis
nur nom.	0,317***			
3. Preis	0,745***	0,428***		
2. Preis	1,576***	1,259*	0,831	
1. Preis	2,365***	2,048***	1,620*	0,789**

* $p < 0.05$, ** $p < 0.025$, *** $p < 0.01$. Abgebildet sind Mittelwertdifferenzen.

Signifikanz der Mittelwertunterschiede mit Rangsummentest nach Kruskal-Wallis ermittelt.

Anhang 2: Abbildungen

Abbildung 1: Typen von Zitationszyklen

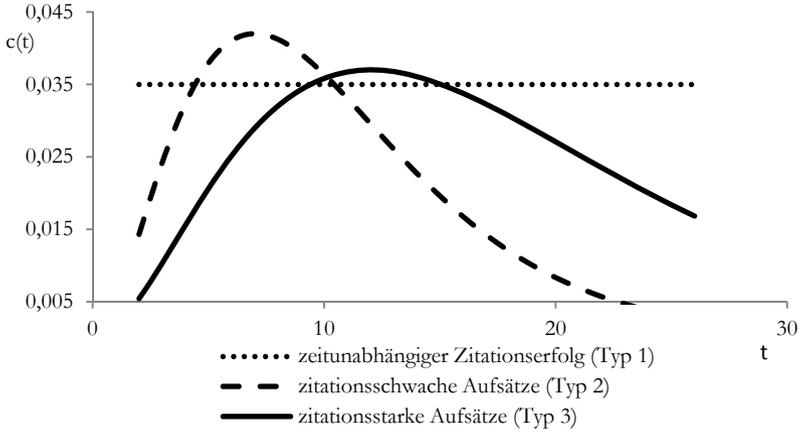


Abbildung 2: Verteilung der (adj.) Zitationen pro Jahr nach Prämierungsstatus

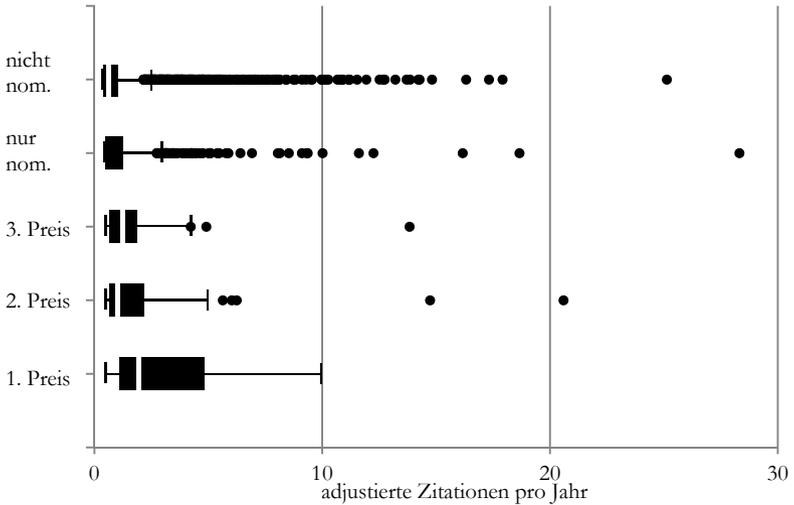
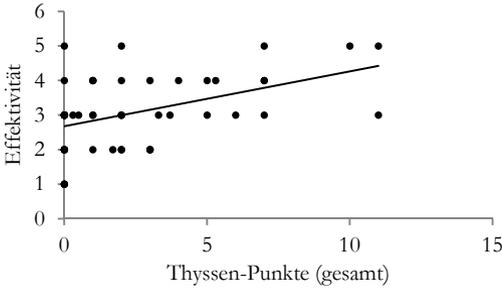
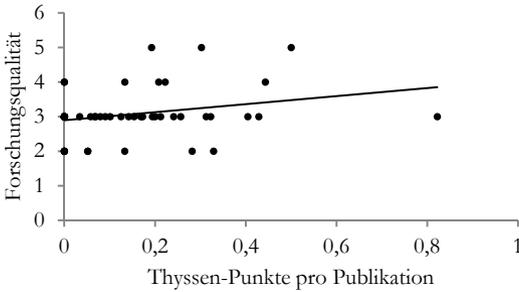


Abbildung 3: Preis der Fritz Thyssen Stiftung und Leistungsindikatoren des Wissenschaftsrates

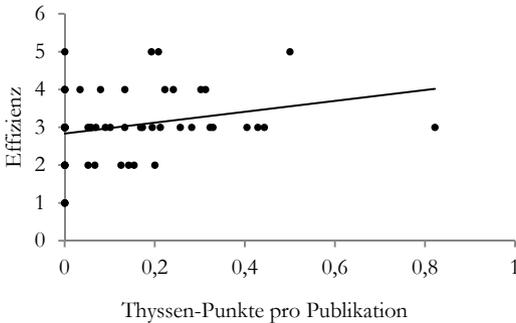
a) Absoluter Zshg. der Jahre 2001 bis 2005, $r = 0,461^{***}$ $n = 53$



(b) Absoluter Zshg. der Jahre 2001 bis 2005, $r = 0,275^{**}$ $n = 54$



(c) Relativer Zshg. der Jahre 2001 bis 2005, $r = 0,251^{**}$ $n = 53$

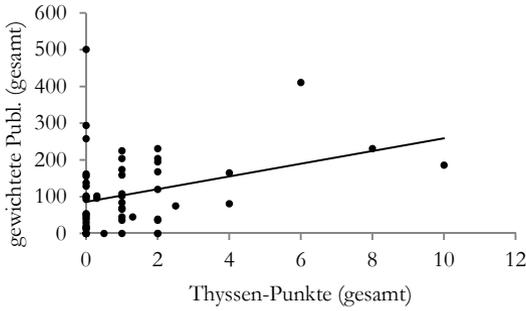


Datenquelle: Leistungsindikatoren des Forschungsratings Soziologie des Wissenschaftsrates (2008), Thyssen-Punkte aus eigenem Datensatz berechnet.

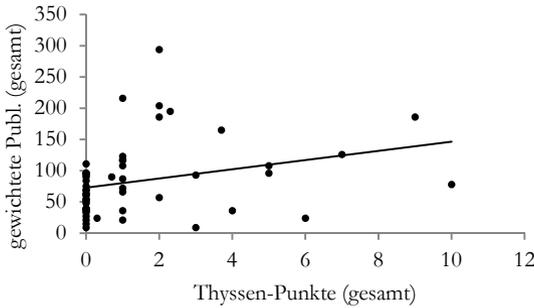
r = Korrelationskoeffizient (nach Pearson); * $p < 0.05$, ** $p < 0.025$, *** $p < 0.01$

Abbildung 4: Preis der Fritz Thyssen Stiftung und Publikationen

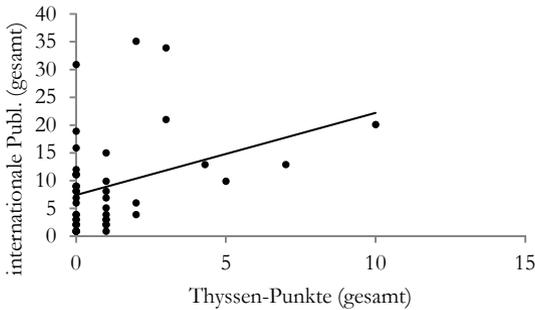
(a) Absoluter Zshg. der Jahre 1998 bis 2000, $r = 0,324^{**}$ $n = 55$



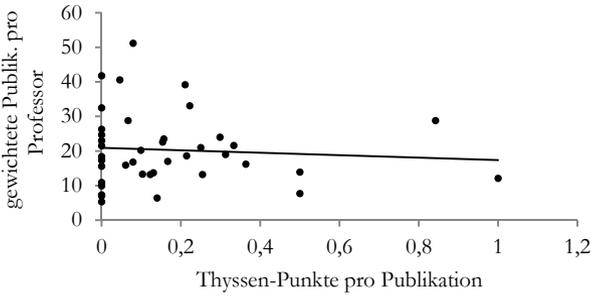
(c) Absoluter Zshg. der Jahre 2001 bis 2003, $r = 0,291^*$ $n = 54$



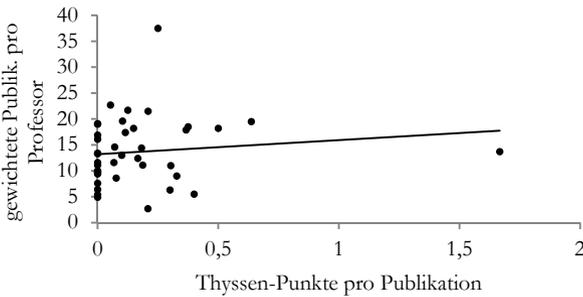
(e) Absoluter Zshg. der Jahre 2004 bis 2006, $r = 0,354^{***}$ $n = 45$



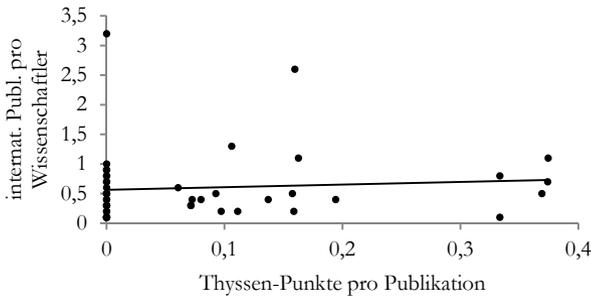
(b) Relativer Zshg. der Jahre 1998 bis 2000, $r = 0,076$ $n = 40$



(d) Relativer Zshg. der Jahre 2001 bis 2003, $r = 0,122$ $n = 41$



(f) Relativer Zshg. der Jahre 2004 bis 2006, $r = 0,087$ $n = 43$



Datenquelle: Publikationszahlen der Hochschulrankings 2002, 2005 und 2008 des CHE, Thysen-Punkte aus eigenem Datensatz berechnet.

r = Korrelationskoeffizient (nach Pearson); * $p < 0.05$, ** $p < 0.025$, *** $p < 0.01$