

Das Handelsblatt Ökonomen-Ranking 2007: Eine kritische Beurteilung

Robert Hofmeister
Universität Konstanz

Heinrich W. Ursprung*
Universität Konstanz

1. Einleitung

Dass Forschungsevaluationen unverzichtbare Führungsinstrumente darstellen, ist heutzutage praktisch unbestritten; zu kleineren Irritationen führt allenfalls noch der Umstand, dass mancherorts zu oft und zu wenig koordiniert evaluiert wird.¹ Was direkt betroffene Wissenschaftler jedoch immer wieder in Frage stellen, ist die *Qualität* von Evaluationen.² Da schlecht konzipierte Evaluationen nicht bloß ein Ärgernis für die direkt Betroffenen darstellen, sondern dem Wissenschaftssystem auch erheblichen Schaden zufügen können, hat der Verein für Socialpolitik vor zwei Jahren eine *Arbeitsgruppe Forschungsmonitoring* (AGFM) ins Leben gerufen. Die wichtigste Aufgabe der AGFM ist es, eine Forschungsdatenbank zu pflegen, mit deren Hilfe Forschungsevaluationen im Bereich der Wirtschaftswissenschaften überprüft und informiert kommentiert werden können.³

Anlässlich der Bayreuther Jahrestagung des *Vereins für Socialpolitik* haben wir in einer Paneldiskussion das *Handelsblatt Ökonomen-Ranking 2006* kommentiert. Wir attestierten damals den Verantwortlichen, das mit Abstand be-

*Fachbereich Wirtschaftswissenschaften, Universität Konstanz, Fach D-138, 78457 Konstanz, Email des korrespondierenden Autors: Heinrich.Ursprung@uni-konstanz.de. – Der Zweitautor dankt dem *Center for Study of Public Choice*, George Mason University, für die Gastfreundschaft.

1. Siehe z.B. Frey (2007).

2. Siehe z.B. Ursprung (2003).

3. Ebenfalls beunruhigt über die mangelnde Qualität breit publizierter Forschungsevaluationen hat der *Wissenschaftsrat* ein eigenes Forschungsrating von Universitäten und außeruniversitären Forschungseinrichtungen initiiert. Nach erfolgreichem Abschluss einer Pilotstudie sollen in einem rotierenden Turnus alle Fachgebiete abgedeckt werden. Zur ursprünglichen Empfehlung siehe <http://www.wissenschaftsrat.de/texte/6285-04.pdf> und zur laufenden Pilotstudie http://www.wissenschaftsrat.de/pilot_start.htm.

Das Handelsblatt Ökonomen-Ranking 2007: Eine kritische Beurteilung

ste privat (d.h. nicht im Wissenschaftssystem) erstellte Forschungsranking für die Volkswirtschaftslehre publiziert zu haben.⁴ In der Zwischenzeit hat das *Handelsblatt* mit dem *Thurgauer Wirtschaftsinstitut*, das die Datenbank der AGFM verwaltet, eine Kooperation vereinbart. Diese Kooperation beinhaltet die Überprüfung ausgewählter VWL-Daten des *Handelsblatts* und die Erstellung eines äquivalenten Datensatzes für die Betriebswirtschaftslehre. Das *Handelsblatt Ökonomen-Ranking 2007* beruht also auf Publikationsdaten, die mit dem Datensatz der AGFM abgeglichen worden sind. Das erste *Handelsblatt Betriebswirte-Ranking* wird mit der Zeit dieselbe Datenqualität aufweisen.

Die enorme Sichtbarkeit, verbunden mit der großen Akzeptanz bei den evaluierten Ökonomen (und vielleicht auch die Zusammenarbeit mit der AGFM) verleihen dem *Handelsblatt-Ranking* einen Stellenwert, der über die offenbar vorhandene „Newsworthiness“ hinausgeht. Das *Handelsblatt-Ranking* wird insbesondere auch innerhalb des Wissenschaftssystems rezipiert und übt so einen nicht zu unterschätzenden Einfluss auf die Profession aus. Gerade weil dem so ist, erscheint es uns wichtig, an dieser Stelle zu zeigen, dass das *Handelsblatt-Ranking*, trotz seiner anderweitigen Verdienste, kein geeignetes Führungsinstrument darstellt. Dabei wollen wir nicht auf die offensichtlichen Grenzen bibliometrischer Indikatoren eingehen, auf die das *Handelsblatt* ja auch immer wieder hinweist. Wir wollen in Abschnitt 2 vielmehr auf drei methodische Eigenheiten des *Handelsblatt-Rankings* hinweisen, die dysfunktionale Anreize zur Folge haben könnten, wenn der zentrale *Handelsblatt*-Indikator zum allgemeinen Evaluationsstandard erhoben würde. In Abschnitt 3 präsentieren wir den von der AGFM präferierten Indikator der Forschungsproduktivität und vergleichen dann in Abschnitt 4 die beiden Methoden anhand des so genannten Lebenswerk-Ranking. Dabei zeigt sich, dass die Evaluationsmethode des *Handelsblatts* zu verzerrten Messungen der Forschungsproduktivität führt, d.h. die Produktivität gewisser Forschertypen wird systematisch über- bzw. unterschätzt.

2. Zur Anreizinkompatibilität des Handelsblatt Ökonomen-Rankings

2.1 Die Zeitschriftenauswahl

Das *Handelsblatt Ökonomen-Ranking 2007* setzt Forschungsoutput mit Artikeln gleich, die in einer Auswahl von 220 besonders angesehenen Fachzeitschriften erschienen oder zur Publikation angenommen worden sind. Das *Handelsblatt* erfasst also nur eine relativ kleine *Auswahl* von einschlägigen wissenschaftlichen Zeitschriften. Nun führt aber jede noch so wohlüberlegte Auswahl leicht zu Verhaltensänderungen, die weder beabsichtigt noch erstrebenswert sind.

4. Eine Kritik am ersten *Handelsblatt Ökonomen-Ranking* vom Herbst 2005, das mit einer anderen Methode erstellt wurde, findet sich in Zimmer und Ursprung (2007).

Robert Hofmeister und Heinrich W. Ursprung

Unschön ist also nicht nur der Umstand, dass der Forschungsoutput unvollständig und deshalb unter Umständen verzerrt erfasst wird; mit einer derartigen Auswahl nimmt man auch gewissen Ökonomen die Anreize, sich mit Problemen auseinanderzusetzen, die z.B. nur von lokalem Interesse sind und somit bloß in „minor journals“ publiziert werden können. Darüber hinaus werden auch all diejenigen „journeymen“ Forscher demotiviert, die das Handwerk der ökonomischen Forschung zwar gut beherrschen, deren Arbeit aber nicht die von renommierten Zeitschriften verlangte hohe Originalität aufweist. Ein Indikator, der aber nur einen speziellen Teil der ökonomischen Forschung erfasst, kann nicht der ganzen Profession als Messlatte dienen.

Ein allgemein einsetzbarer Forschungsindikator muss notwendigerweise eine Vollerhebung des Forschungsoutputs anstreben. Dieses Ziel ist zugegebenermaßen nie wirklich erreichbar, da die Grenze zwischen Forschung und Anwendung wissenschaftlicher Methoden fließend ist und mit rein bibliometrischen Mitteln ohnehin nicht präzise identifiziert werden kann. Das Kriterium der Begutachtung durch anonyme Sachverständige ist jedoch wenig kontrovers und hat sich auch deshalb etabliert, weil es bibliometrisch relativ leicht umsetzbar ist. Für die Ökonomie bedeutet dies, dass man Forschung mit Publikationen in ökonomischen Fachzeitschriften mit Begutachtungsverfahren gleichsetzen kann; und da die von *EconLit* auferlegten Minimalstandards dem Begutachtungs-Kriterium recht nahe kommen, hat sich in der Bibliometrie die von *EconLit* erfasste Zeitschriftenliteratur als Standard durchgesetzt.

Die über 750 Zeitschriften, die heutzutage von *EconLit* erfasst werden, umfassen praktisch den gesamten Kernbereich der ökonomischen Forschung. Für gewisse Subdisziplinen, die an den Schnittstellen der Ökonomie und einer anderen Disziplin angesiedelt sind, ist der *EconLit*-Standard aber nicht befriedigend. Um den Forschungsoutput von Ökonometrikern besser abzubilden zu können, erfasst das *Handelsblatt* deshalb neuerdings auch Zeitschriften im Bereich der Statistik. Neben der Ökonometrie gibt es aber noch weitere Subdisziplinen, in denen häufig in nicht-ökonomischen Zeitschriften publiziert wird. Zu nennen sind hier z.B. die Politische Ökonomie, die Wirtschaftsgeschichte und die Verhaltensökonomik.

2.2 Die Koautorengewichtung

Die vom *Handelsblatt* verwendete Koautorengewichtung, die jedem der n Autoren eines Artikels den Anteil $2/(n + 1)$ gutschreibt, ist aus verschiedenen Gründen problematisch. Im Vordergrund steht hier die mit dieser Gewichtung einhergehende Anreizinkompatibilität. Wenn ein von vielen Autoren verfasster Artikel bis zu zweimal mehr Gewicht erhält als ein von einem einzigen Autor verfasster Artikel, so werden Anreize für Koautorschaften geschaffen, ob dies im gegebenen Fall nun sinnvoll ist oder nicht. Im Extremfall kann diese Gewichtung sogar zu fingierten Koautorschaften führen, die im Logrollingverfahren abgesprochen werden, oder zu Koautorschaften ganzer Forschungseinheiten im Stil naturwissenschaftlicher Labors. Die Standardmethode der

Das Handelsblatt Ökonomen-Ranking 2007: Eine kritische Beurteilung

Koautorengewichtung ($1/n$ anstatt $2/(n + 1)$) weist keine derartigen dysfunktionalen Anreize auf. Darüber hinaus trägt die Standardmethode auch dem Umstand Rechnung, dass in den verschiedenen Subdisziplinen der Ökonomie womöglich eine unterschiedliche Anzahl von Autoren optimal und üblich ist: Wenn es nämlich derartige Unterschiede gibt, so zieht die Anrechnungsmethode des *Handelsblatts* systematische Verzerrungen nach sich.⁵

2.3 Die (vernachlässigte) Gewichtung des Umfangs von Forschungsergebnissen

Man mag unsere Befürchtung, dass die Koautorengewichtung des *Handelsblatts* missbraucht werden könnte, vielleicht für übertrieben halten. Unser letzter Einwand, der sich gegen die Vernachlässigung jeglicher Gewichtung des Umfangs der erfassten Zeitschriftenartikel wendet, stützt sich jedoch auf ein Missbehagen, das in der Profession weit verbreitet ist. Wenn die (qualitätsgewichtete) Anzahl der Publikationen ganz allgemein als valider Indikator der Forschungsproduktivität interpretiert wird, so verleitet dies dazu, den Forschungsoutput in „kleinste publizierbare Einheiten“ aufzuteilen. Abgesehen davon, dass eine derartige Fraktionierung die Kommunikation und Rezeption von Forschungsergebnissen behindert und den Begutachtungsprozess der Fachzeitschriften unnötig belastet, führt diese Bemessungsgrundlage auch zu verzerrten Messungen, weil eine Fraktionierung der Forschungsergebnisse nicht in allen Subdisziplinen im gleichen Umfang möglich ist.

Um dem Unwesen der Balkanisierung der Forschungsergebnisse zu begegnen, bietet es sich an, bei der Messung des Forschungsoutputs die Seitenzahl der Artikel zu berücksichtigen. Die Seitenzahl kann relativ einfach erhoben werden und reflektiert, weil der Zeitschriftenplatz beschränkt ist, die Einschätzung der Bedeutung der jeweiligen Artikel in den Augen der Herausgeber.⁶ Diese Längenunterschiede sind nicht unbeträchtlich. So sind z.B. fast ein Drittel aller AER-Artikel, die von in Deutschland tätigen Ökonomen verfasst wurden, relativ kurz (bis zu 10 Seiten), währenddem doch etwa 20% mehr als 20 Seiten aufweisen.

Obwohl die meisten bibliometrischen Studien mit einer Seitenzahl-Gewichtung arbeiten, sollte diese Art der Gewichtung aber bloß als grober Indikator der Bedeutung betrachtet werden. Verschiedene Verfeinerungen sind denn auch in der bibliometrischen Literatur vorgeschlagen worden. So wird

5. Es ist richtig, dass Artikel, die von mehreren Autoren verfasst worden sind, ceteris paribus eine höhere Qualität aufweisen (siehe z.B. Rauber und Ursprung, 2008). Dies sollte aber nicht in der Koautorengewichtung berücksichtigt werden, da dieser Effekt ja schon im Qualitätsgewicht der publizierenden Zeitschrift zum Ausdruck kommt.
6. Dass diese Einschätzung der Herausgeber durchaus fundiert ist, kann man daraus ersehen, dass längere Zeitschriftenartikel (ceteris paribus) öfter zitiert werden als kürzere. Eine entsprechende Untersuchung findet sich in Zimmer und Ursprung (2007).

z.B. die Anzahl der Schriftzeichen pro Artikel verwendet, um unterschiedliche Seitenlayouts abzubilden, oder die Seitenzahl im Vergleich mit anderen Aufsätzen in der jeweiligen Zeitschrift bzw. Subdisziplin, um zeitschriftenspezifische oder subdisziplinspezifische Gepflogenheiten zu berücksichtigen.⁷

3. Die Standardmethode der Messung des Forschungsausputs

Die Standardmethode der Messung des Forschungsausputs, die mit den im vorhergehenden Abschnitt diskutierten Nachteilen nicht behaftet ist, beruht auf der folgenden Definition des Forschungsausputs F_i von Forscher i :

$$F_i = \sum_{k \in EL} \frac{w_k p_k}{n_k}. \quad (1)$$

Dabei bezeichnet w_k die Qualität der Zeitschrift, in der Artikel k von Forscher i erschienen ist, p_k die Seitenzahl des Artikels und n_k die Anzahl der Autoren. Die Summe umfasst alle von *EconLit* (EL) erfassten Zeitschriftenartikel. Die entsprechende vom *Handelsblatt* verwendete Formel sieht folgendermaßen aus:

$$H_i = \sum_{k \in HB} \frac{2\hat{w}_k}{(n_k + 1)}, \quad (2)$$

wobei die Summe nun über alle Artikel k läuft, die in einer vom *Handelsblatt* (HB) erfassten Zeitschrift veröffentlicht oder zur Veröffentlichung angenommen worden sind.⁸ Man beachte, dass die Zeitschriftenauswahl (HB) des *Handelsblatts* im Vergleich zur gesamten *EconLit* Literatur (EL) relativ klein ist. Die im folgenden Abschnitt betrachteten Ökonomen haben z.B. 8130 Artikel in 527 verschiedenen *EconLit* Zeitschriften veröffentlicht; davon werden jedoch nur 4365 Artikel in 168 Zeitschriften vom *Handelsblatt* erfasst.

Bei einer automatisierten Datenerhebung fallen keine beträchtlichen zusätzlichen Kosten an, wenn man die Methode des *Handelsblatts* durch die von der AGFM präferierten Standardmethode ersetzt. Die Koautorengewichtung kann man praktisch durch Knopfdruck verändern, und der *EconLit* Datensatz enthält

7. Wenn man die Länge der Artikel mit der in der entsprechenden Zeitschrift üblichen Artikellänge vergleicht, so vermeidet man damit auch einen möglichen Fehler. Man kann nämlich einwenden, dass die Länge von Aufsätzen schon in der Qualitätsgewichtung der Zeitschrift enthalten ist, wenn die Qualitätsgewichtung den „impact“ der Zeitschrift widerspiegelt und der „impact“ wiederum von der Länge abhängt.

8. Die Zeitschriftengewichte w_k in Formel 1 hat die AGFM bisher der Studie von Combes und Linzner (2003) entnommen. Das *Handelsblatt* orientiert sich bei seiner Zeitschriftengewichtung \hat{w}_k (Formel 2) ebenfalls weitgehend an dieser Studie. Wir gehen hier nicht auf den offensichtlichen Einfluss der Qualitätsgewichtung der Zeitschriften ein und lassen die Frage der adäquaten Gewichtung offen (siehe dazu auch die Beiträge von Klaus Ritzberger und Günther Schulze et al. in diesem Band).

Das Handelsblatt Ökonomen-Ranking 2007: Eine kritische Beurteilung

alle Informationen über die vom *Handelsblatt* bisher nicht erfassten Zeitschriften und Seitenzahlen. Ein Wermutstropfen besteht allerdings darin, dass man die im Erscheinen begriffenen Arbeiten mit der Methode der AGFM nicht ohne weiteres berücksichtigen kann, weil die entsprechenden Seitenzahlen eben noch nicht bekannt sind.

Im vorhergehenden Abschnitt haben wir darauf hingewiesen, dass man mit der in Formel (2) zusammengefassten Methode des *Handelsblatts* neben den unerwünschten Anreizwirkungen auch Verzerrungen bei der Messung des Forschungsoutputs in Kauf nimmt. Wenn man zeigen könnte, dass diese Verzerrungen geringfügig sind, so fielen die Vorteile der doch etwas aufwändigeren Standardmethode (Formel 1) vielleicht nicht wirklich ins Gewicht. Dies gilt allerdings nur solange man lediglich die gewachsene Forschungslandschaft abbilden will und niemand den Handelsblatt-Indikator als Führungsinstrument missbraucht. Im nächsten Abschnitt vergleichen wir konkrete „Lebenswerk-Rankings“, die mit der Standardmethode und Elementen der Handelsblatt-Methode erstellt wurden, und zeigen, dass die Verzerrungen, die man sich mit der Handelsblatt-Methode einhandelt, doch erheblich sind.

4. Das *Handelsblatt* Ranking auf dem Prüfstand

4.1 Das „Lebenswerk“ Ranking

In Tabelle 1 präsentieren wir das nach der Standardmethode (Formel 1) erstellte Lebenswerk-Ranking der an deutschen Universitäten tätigen Volkswirte, deren Lebenswerk gemäß dem von Combes und Linnemer vorgeschlagenen Gewichtungsschema mehr als 100 AER-äquivalente Seiten aufweist ($F_i > 100$). Man beachte, dass dieses Ranking im Vergleich mit dem *Handelsblatt-Ranking*, das ja auch akzeptierte, aber noch nicht publizierte Artikel erfasst, etwas weniger aktuell ist. Wir ziehen das Lebenswerk-Ranking zur Illustration der methodischen Unterschiede heran, weil das *Handelsblatt* mit diesem Ranking seine „top-Ökonomen“ identifiziert und nicht etwa weil wir der Ansicht wären, dass derartige Lebenswerk-Rankings wirklich eine große Bedeutung hätten. Das Lebenswerk ist bestenfalls ein Maß für den Erfahrungsschatz eines Wissenschaftlers. Wenn man aus dem Lebenswerk Rückschlüsse auf die Leistungsfähigkeit ziehen will, so muss man entweder das Karrierealter in Rechnung stellen, so wie dies das *Handelsblatt* in seinem „Punkte pro Jahr“ Ranking tut oder, besser noch, den Leistungsvergleich auf eine Alterskohorte beschränken, um den Einfluss des Alters adäquat zu berücksichtigen (siehe dazu Rauber und Ursprung, 2008).

Die von der AGFM präferierte Standardmethode (Formel 1) dient unserer Analyse als Vergleichsbasis. Davon ausgehend untersuchen wir den Einfluss der drei monierten Eigenheiten der *Handelsblatt*-Methode. Wir vergleichen also nicht das vom *Handelsblatt* publizierte Lebenswerk-Ranking mit alternativen Rankings. Wir nehmen vielmehr das in Tabelle 1 (trunkiert) präsentierte „Standardranking“ und zeigen wie es sich verändert, wenn man nacheinander

Robert Hofmeister und Heinrich W. Ursprung

Tabelle 1 „Lebenswerk“ Ranking

Rang	Name	Lebenswerk -Output F	Rang	Name	Lebenswerk -Output F
1	Hellwig, Martin	F = 501,54	29	Funke, Michael	128,78
2	Sinn, Hans-Werner	362,01	30	Rees, Ray	128,10
3	Güth, Werner	257,53	31	Franz, Wolfgang	127,56
4	Konrad, Kai A.	255,15	32	Egger, Peter	125,79
5	von Hagen, Jürgen	239,94	33	Ursprung, Heinrich W.	124,31
6	Börsch-Supan, Axel	230,97	34	Janeba, Eckhard	123,93
7	Schmidt, Klaus	221,67	35	Belke, Ansgar	123,67
8	Ebert, Udo	212,88	36	Oechssler, Jörg	123,36
9	Bester, Helmut	210,48	37	Strausz, Roland	123,25
10	Vaubel, Roland	207,63	38	Wieland, Volker	119,91
11	Snowder, Dennis	201,61	39	Pethig, Rüdiger	116,06
12	Moldovanu, Benny	188,83	40	Dawid, Herbert	115,13
13	Backhaus, Jürgen G.	176,14	41	Eichberger, Jürgen	114,07
14	Schweizer, Urs	168,14	42	Walz, Uwe	112,04
15	Bell, Clive	159,82	43	Schnitzer, Monika	111,67
16	Kraft, Kornelius	155,94	44	Röller, Lars-Hendrik	111,61
17	Komlos, John	155,04	45	Härdle*, Wolfgang	109,78
18	Gerlach, Stefan	150,22	46	Peitz, Martin	108,76
19	von Thadden, E.-L.	148,33	47	Buch, Claudia	108,38
20	Burda, Michael C.	143,54	48	Falk, Armin	106,97
21	Corneo, Giacomo	142,46	49	Riphahn, Regina T.	106,88
22	Zimmermann, Klaus F.	141,81	50	Puppe, Clemens	105,75
23	Endres, Alfred	139,26	51	Fuest, Clemens	105,35
24	Witt, Ulrich	137,06	52	Breyer, Friedrich	103,88
25	Schmitz, Patrick	134,97	53	Welsch, Heinz	101,24
26	Feld, Lars P.	134,10	54	Bolle, Friedel	100,97
27	Wagner, Joachim	129,76	55	Leininger, Wolfgang	100,76
28	Krebs, Tom	129,42	56	Klasen, Stephan	100,54

Quelle: Eigene Berechnungen.

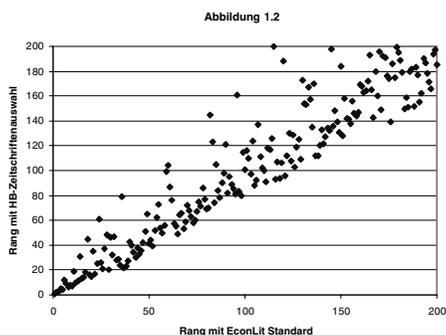
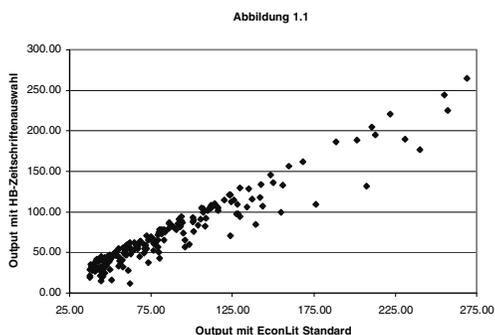
*Wolfgang Härdle ist der einzige Ökonometriker, der auf dieser Liste auftaucht. Er und seine Kollegen werden von *EconLit* basierten Rankings unterbewertet.

die Gewichtungen der Standardmethode durch die monierten Gewichtungen der *Handelsblatt*-Methode ersetzt. Der Vorteil dieses Vorgehens besteht darin, dass wir dazu nur den AGFM Datensatz benötigen.

In einem ersten Schritt beschränken wir die Publikationserfassung auf diejenigen *EconLit*-Zeitschriften, die vom *Handelsblatt* ausgezählt werden.⁹ Wir versehen also alle *EconLit*-Zeitschriften, die nicht in der *Handelsblatt*-Auswahl enthalten sind, mit dem Gewicht 0.

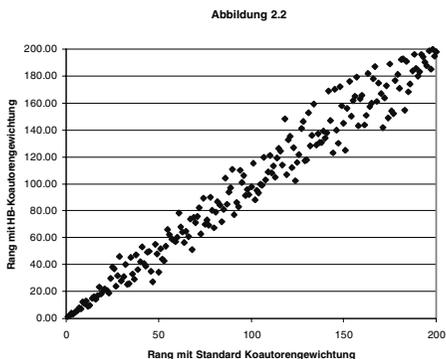
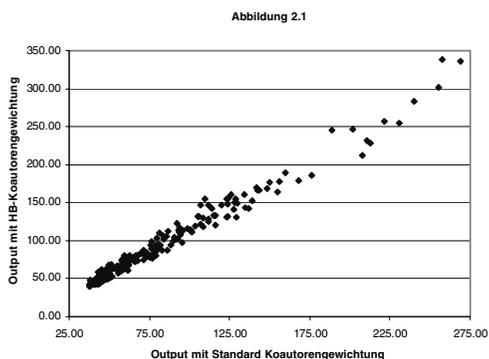
9. Man beachte, dass wir die HB-Auswahl nicht genau reproduzieren können, weil die Datenbank der AGFM vorderhand nur die *EconLit*-Zeitschriften umfasst.

Das Handelsblatt Ökonomen-Ranking 2007: Eine kritische Beurteilung



Quelle: Eigene Berechnungen.

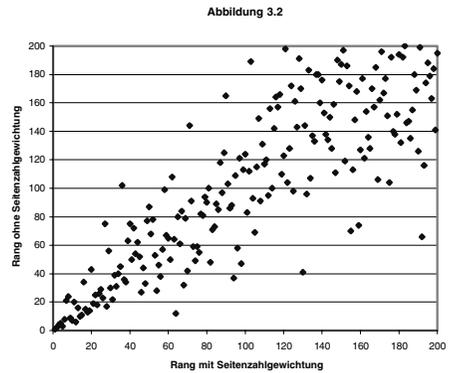
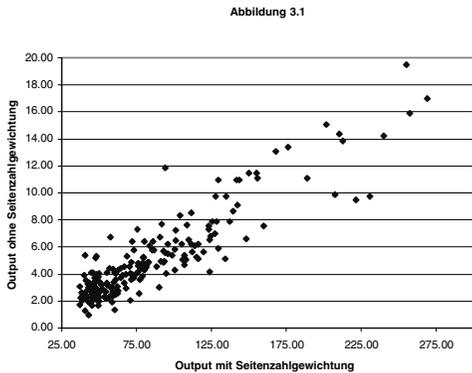
Die Streudiagramme 1.1 und 1.2 zeigen wie sich beim Übergang von der „EconLit Vollerhebung“ zur HB-Auswahl der gemessene Forschungsoutput und die sich daraus ergebenden Rangplätze ändern. Die ganz erheblichen Rangverschiebungen im Vergleich zur Vollerhebung weisen darauf hin, dass die Stichprobe des *Handelsblatts* in der Tat Verzerrungen zur Folge hat. Die Beschränkung auf die *Handelsblatt*-Auswahl führt zu individuellen Rangpunkteverbesserungen von bis zu 20% und zu noch sehr viel größeren Einbußen. Darüber hinaus zeigt sich, dass eben auch international durchaus sichtbare deutsche Volkswirte (alle Streudiagramme beschränken sich auf die top-200 Ökonomen des Standard-Ranking)¹⁰ einen erheblichen Teil ihres Oeuvres in „minor journals“ publizieren.



Quelle: Eigene Berechnungen.

Im zweiten Schritt analysieren wir den Einfluss der Koautorengewichtung. Die Streudiagramme 2.1. und 2.2 dokumentieren wie sich die doch sehr ei-

10. Die beiden führenden Ökonomen fallen allerdings bei den Output-Abbildungen „aus dem Rahmen“.



Quelle: Eigene Berechnungen.

genwillige Koautorengewichtung des *Handelsblatts* auf den gemessenen Forschungsausput und die Rangierung der „Top-Ökonomen“ auswirkt. Die Verzerrungen sind hier geringer als beim ersten Vergleich, aber immer noch durchaus beträchtlich. Eine genauere Betrachtung der Datenpunkte bestärkt uns in der Vermutung, dass z.B. experimentelle Ökonomen, die regelmäßig mit mehreren Koautoren zusammenarbeiten, von der *Handelsblatt*-Methode profitieren, wohingegen Ökonomen, die unorthodoxe Forschungsstrategien verfolgen und somit eher als wissenschaftliche Einzelkämpfer auftreten, von der *Handelsblatt*-Methode unterbewertet werden.

Die markantesten Änderungen ergeben sich jedoch bei unserem letzten Vergleich, in dem wir wie das *Handelsblatt* auf jegliche Gewichtung des Umfangs der Forschungsergebnisse verzichten, d.h. wir geben die Seitenzahl-Gewichtung auf. Die Bedeutung der Seitenzahlengewichtung kommt in den Streudiagrammen 3.1 und 3.2 sehr gut zum Ausdruck. Es gibt offenbar Ökonomen, die regelmäßig Forschungsergebnisse publizieren, deren Präsentation die volle Länge eines Zeitschriftenartikels verlangt und rechtfertigt, währenddem andere sich auf punktuelle Probleme spezialisieren, deren Gesamtzusammenhang wohlbekannt ist und die deshalb relativ kurz abgehandelt werden können.

4.2 Fachbereichsrangings

Die Eigenheiten bibliometrischer Evaluationsmethoden machen sich natürlich auch im Aggregat, d.h. bei der Evaluation gesamter Fachbereiche bemerkbar. Tabelle 2 vergleicht das vom *Handelsblatt* an erster Stelle aufgeführte Fachbereichsrangung, dem der gesamte Forschungsausput (seit 1997) der gegenwärtigen Mitarbeiter zugrunde liegt, mit dem entsprechenden Rangung, das mit der Standardmethode erstellt wurde. Wir beschränken uns dabei auf eine top-15 Rangliste der deutschen Fachbereiche. Weil der gesamte Forschungsausput eines Fachbereichs natürlich ganz wesentlich von der Anzahl der ihm

Das Handelsblatt Ökonomen-Ranking 2007: Eine kritische Beurteilung

Tabelle 2 Evaluation von Fachbereichen

Gesamtoutput		Durchschnittsproduktivität der Professoren	
<i>Handelsblatt</i>	Standardmethode	<i>Handelsblatt</i>	Standardmethode
1 Bonn	Bonn	Bonn	Bonn
2 LMU	LMU	Mannheim	LMU
3 Mannheim	Mannheim	Frankfurt/M	Oldenburg
4 Frankfurt/M	Frankfurt/M	Kiel	Mannheim
5 HU Berlin	FU Berlin	LMU	Hannover
6 Köln	Hannover	HU Berlin	Heidelberg
7 FU Berlin	HU Berlin	Heidelberg	Kiel
8 Kiel	Kiel	Bielefeld	Tübingen
9 Bielefeld	Köln	Dortmund	Frankfurt/M
10 Hannover	Heidelberg	Konstanz	Konstanz
11 Heidelberg	Bielefeld	Oldenburg	FU Berlin
12 Konstanz	Hamburg	Köln	Bielefeld
13 Dortmund	Tübingen	FU Berlin	HU Berlin
14 Hamburg	Konstanz	Hannover	Marburg
15 Karlsruhe	Erfurt	Tübingen	Frankfurt/O

angehörigen Mitarbeiter abhängt, birgt dieses Ranking-Kriterium die Gefahr in sich, falsch interpretiert zu werden. In mancherlei Hinsicht aufschlussreicher sind Produktivitätsmaße, die eben nicht direkt von der Größe der Fachbereiche abhängen. Wir reproduzieren deshalb auch das Produktivitätsranking des *Handelsblatts* (Punkte pro Professor) und stellen es dem entsprechenden Ranking gegenüber, das mit der Standardmethode erstellt wurde. Während beim Ranking nach dem Gesamtoutput die Rangordnung nicht allzu stark von der verwendeten Methode abhängt, ergeben sich beim Produktivitätsranking doch große Unterschiede.

Nicht überraschend ist, dass sich die Ranglisten nach Gesamtoutput und Forschungsproduktivität stark unterscheiden. Es ist wohlbekannt, dass man umfassende Evaluationen nicht auf ein einziges Kriterium abstellen kann. Obwohl der „default mode“ der *Handelsblatt*-Homepage die Fachbereiche nach dem problematischen Kriterium „Gesamtoutput“ auflistet, weist das *Handelsblatt* in vorbildlicher Weise weitere Indikatoren aus. Neben der schon angeführten Forschungsproduktivität geben diese zusätzlichen Indikatoren Hinweise auf die Qualität der Forschung (top-5 Punkte) und auf die Streuung der Forschungsproduktivität innerhalb der jeweiligen Fachbereiche (Anteil des forschungsstärksten Professors an der Gesamtpunktezahl). Eine derartige mehrdimensionale Evaluation ist sehr nützlich und sinnvoll, weil die Interessen der Nutzer von Rankings eben sehr unterschiedlich sein können.

Obwohl gerade die Mehrdimensionalität des *Handelsblatt*-Fachbereichsranking lobend zu erwähnen ist, lassen sich die Aussagekraft des Qualitäts- und Streuungsindikators und die Präsentation der facettenreichen Information sicherlich noch verbessern. Wir schlagen in Abbildung 4 eine

Robert Hofmeister und Heinrich W. Ursprung

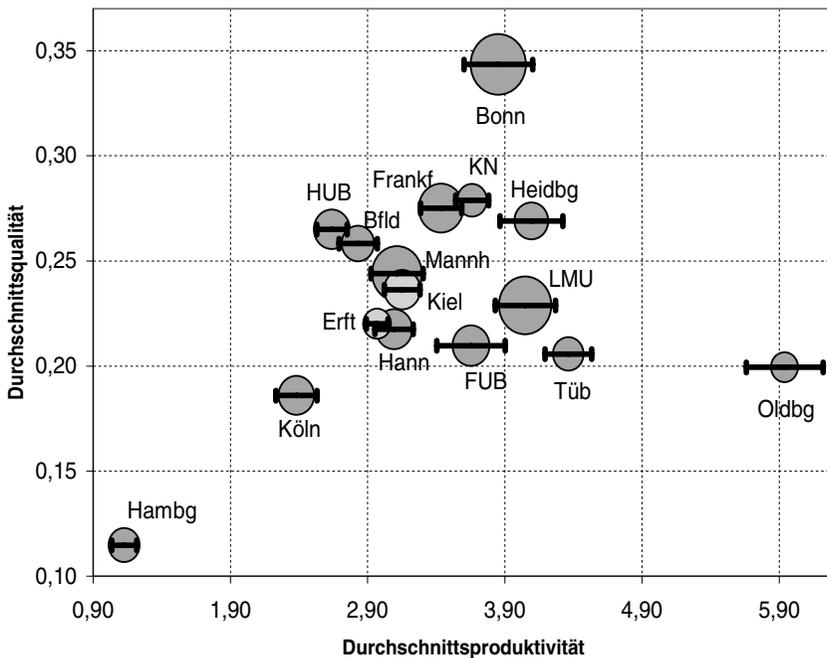


Abbildung 4

Präsentationsform vor, in der die vielleicht vier wichtigsten Indikatoren gleichberechtigt nebeneinander stehen und zudem auf einen Blick erfasst werden können. Alle vier Indikatoren wurden mit Hilfe der Standardmethode berechnet und beziehen sich auf alle Mitarbeiter, d.h. es wird auch bei der Berechnung der Produktivität (und somit anders als im Produktivitätsranking in Tabelle 2) der Forschungsoutput aller promovierten Mitarbeiter erfasst. Auf der Abszisse ist die durchschnittliche Produktivität der einem Fachbereich angehörenden Mitarbeiter abgetragen, auf der Ordinate die Durchschnittsqualität des Forschungsoutputs. Die Fläche der „Fachbereichsbubbles“ entspricht dem Gesamtoutput (seit 1997) und die Länge der Balken ist proportional zur Standardabweichung der Forschungsproduktivitäten der einem Fachbereich angehörenden Mitarbeiter. Es werden die 15 nach dem Gesamtoutput größten Fachbereiche abgebildet.

Aus Abbildung 4 kann man z.B. entnehmen, dass die Fachbereichsgröße keinen Einfluss auf die Effizienz der Forschungsproduktion zu haben scheint. Sowohl der große Fachbereich der Universität Bonn, wie auch der mittelgroße Fachbereich der Universität Heidelberg und der kleine Fachbereich der Universität Oldenburg produzieren an der offenbaren Effizienzgrenze von Produktivität und Qualität. Die Belegschaft dieser Fachbereiche ist allerdings sehr heterogen. Homogener sind z.B. die ebenfalls recht effizienten Fachbereiche der Universitäten Frankfurt a.M. und Konstanz.

Das Handelsblatt Ökonomen-Ranking 2007: Eine kritische Beurteilung

5. Schluss

Ziel unserer Ausführungen war es zu zeigen, dass das in seiner Art ausgezeichnete *Handelsblatt Ökonomen-Ranking 2007* nicht wirklich anreizkompatibel ist und somit nicht als Führungsinstrument gebraucht werden sollte. Die vom *Handelsblatt* in den letzten zwei Jahren verwendete Methode führt darüber hinaus auch zu erheblich verzerrten Messungen der Forschungsproduktivität. Wir belegen diese Verzerrungen indem wir zeigen, wie sich ein konkretes Ranking ändert, wenn man die bibliometrische Standardmethode, die von der *Arbeitsgruppe Forschungsmonitoring* des *Vereins für Socialpolitik* präferiert wird, durch Elemente der *Handelsblatt*-Methode ersetzt. Es ging uns hier aber nicht darum, einen eigenen VfS-Standard zu entwerfen bzw. zu vertreten. Wir hoffen allerdings, dass ein entsprechender Diskurs mit der Zeit dazu führt, dass sich ein derartiger VfS-Standard herauskristallisiert.

Abschließend kann man festhalten, dass es dem *Handelsblatt* gelungen ist, das wegweisende Ranking aus dem Jahr 2006 weiter zu verbessern. Das neueste *Handelsblatt Ökonomen-Ranking* stützt sich auf verlässlichere Daten, ist differenzierter angelegt und wird einer der interdisziplinär ausgerichteten Subdisziplinen, nämlich der Ökonometrie, besser gerecht. Die Profession darf sich glücklich schätzen, dass ihre Forschungsleistungen von einem derart kompetenten privaten Anbieter einer breiten Öffentlichkeit vorgestellt werden. Wir haben hier auf die Grenzen dieses Rankings hingewiesen und Verbesserungsmöglichkeiten aufgezeigt, die es natürlich immer gibt. Wie das *Handelsblatt* darauf reagieren will, bleibt ihm überlassen. Die Profession der Volkswirte wartet auf alle Fälle gespannt auf das *Handelsblatt Ökonomen-Ranking 2008*.

Literaturverzeichnis

- Combes, P. und L. Linnemer (2003), Where Are the Economists Who Publish? Publication Concentration and Rankings in Europe Based on Cumulative Publications, *Journal of the European Economic Association* 1, 1250–1308.
- Frey, B.S. (2007), Evaluierungen, Evaluierungen, Evaluitis, *Perspektiven der Wirtschaftspolitik* 8, 207–220.
- Rauber, M. und H.W. Ursprung (2008), Evaluation of Researchers: A Life Cycle Analysis of German Academic Economists, in: M. Albert, D. Schmidtchen und S. Voigt (Hrsg.), *Scientific Competition. Conferences on New Political Economy* 25. Mohr Siebeck, Tübingen.
- Ritzberger, K. (2008), Eine invariante Bewertung wirtschaftswissenschaftlicher Fachzeitschriften, *Perspektiven der Wirtschaftspolitik*, dieses Heft.
- Schulze, G., S. Warning und C. Wiermann (2008), Zeitschriftenrankings für die Wirtschaftswissenschaften: Konstruktion eines umfassenden Metaindexes, *Perspektiven der Wirtschaftspolitik*, dieses Heft.
- Ursprung, H.W. (2003), Schneewittchen im Land der Klapperschlangen: Evaluation eines Evaluators, *Perspektiven der Wirtschaftspolitik* 2, 177–190.

Robert Hofmeister und Heinrich W. Ursprung

Zimmer, M. und H.W. Ursprung (2007), Who Is the "Platzhirsch" of the German Economics Profession? A Citation Analysis, *Jahrbücher für Nationalökonomie und Statistik* 227, 187–208.

Abstract: The *Handelsblatt* Ökonomen-Ranking is as yet the best privately provided ranking of individual economists and entire economics departments in Austria, Germany and Switzerland. It is, however, not suited for management purposes because the employed evaluation method gives rise to dysfunctional incentives. The research productivities derived by the *Handelsblatt* method, moreover, turn out to be biased. The *Committee for Research Monitoring* of the *Verein für Socialpolitik* makes use of an established bibliometric method that is not afflicted with these shortcomings. We show how selected results are altered when instead of the standard method various aspects of the *Handelsblatt* method are applied.