

Paludkiewicz, Karol; Wohlrabe, Klaus

## Article

# Qualitätsanalyse von Zeitschriften in den Wirtschaftswissenschaften – über Zitationsdatenbanken und Impaktfaktoren im Online-Zeitalter

ifo Schnelldienst

## Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Paludkiewicz, Karol; Wohlrabe, Klaus (2010) : Qualitätsanalyse von Zeitschriften in den Wirtschaftswissenschaften – über Zitationsdatenbanken und Impaktfaktoren im Online-Zeitalter, ifo Schnelldienst, ISSN 0018-974X, ifo Institut für Wirtschaftsforschung an der Universität München, München, Vol. 63, Iss. 21, pp. 18-28

This Version is available at:

<https://hdl.handle.net/10419/164875>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Qualitätsanalyse von Zeitschriften in den Wirtschaftswissenschaften – über Zitationsdatenbanken und Impaktfaktoren im Online-Zeitalter

18

Karol Paludkiewicz und Klaus Wohlrabe

Impaktfaktoren spielen sowohl in der Wissenschaft als auch der öffentlichen Wahrnehmung immer noch eine große Rolle, wenn es um die Bewertung von Publikationen von Wissenschaftlern geht. Die bekanntesten Impaktfaktoren gehen auf das Web of Science zurück, sind jedoch im Bereich der Wirtschafts- und Sozialwissenschaften einiger Kritik ausgesetzt. Basis der Impaktanalyse sind die wechselseitigen Zitate zwischen den Zeitschriften. Der vorliegende Artikel beschreibt die Funktionsweise sowohl des Web of Science als auch dessen Alternativen Google Scholar und das RePEc-Netzwerk als Zitationsdatenbanken. Letzteres ist auf die Wirtschaftswissenschaften spezialisiert. Darüber hinaus werden Hinweise gegeben, inwieweit infometrische Kennzahlen berechnet und interpretiert werden können. Anhand mehrerer Beispiele werden die Vor- und Nachteile der einzelnen Datenbanken aufgezeigt.

Die (quantitative) Bewertung wissenschaftlicher Forschung spielt eine immer größere Rolle sowohl in der öffentlichen Berichterstattung als auch in der akademischen Beurteilung von Institutionen und Wissenschaftlern. Zentraler Punkt dieser Einschätzung ist meist die Anzahl der Veröffentlichungen in wissenschaftlichen Zeitschriften. Da diese Anzahl in den letzten Jahren stark gestiegen ist, spielt die Qualitätseinschätzung von Zeitschriften eine große Rolle. Sie beruht vor allem auf so genannten Impaktfaktoren, welche wiederum auf den Zitierungen anderer Publikationen (Zeitschriften, Büchern, oder Arbeitspapieren) beruhen. Unter der Annahme, dass Wissenschaftler Arbeiten zitieren, die sie für ihre eigene Forschung als nützlich erachten, ist die Anzahl der Zitate als ein quantitatives Maß für die Resonanz und den Einfluss dieser Publikation innerhalb der Wissenschaft zu sehen. Auf diesen und anderen Qualitätsmerkmalen wird zwischen den verschiedenen Zeitschriften im Hinblick auf Prestige, Einfluss und Stellung in der jeweiligen Fachdisziplin unterschieden. Der Wunsch nach »objektiven« Kriterien zur Unterscheidung spielt in einer Welt, in der die Wissenschaft einen enormen Zuwachs und Vielfalt erfahren hat, eine große Rolle. Zitate und Zeitschriftenrankings besitzen auch eine zunehmende Bedeutung bei der Evaluierung von Forschungsleistungen wirtschaftswissenschaftlicher Fakultäten (vgl. Bommer und Ursprung 1998) und Volkswirten (vgl. Coupe 2003; Ursprung und Zimmer 2007; für eine kritische Sicht der Rankings vgl. Frey und Rost 2010).

Für ökonomische Zeitschriften werden die einflussreichsten Impaktfaktoren auf

Basis der Zitierungen aus dem Web of Science – ein Produkt der Firma Thomson-Scientific – berechnet. Während im Web of Science die Abdeckung für die Naturwissenschaften als gut angesehen wird, wird sie für die Sozialwissenschaften, die in dem Social Science Citation Index (SSCI), zusammengefasst sind, als tendenziell schlecht eingeschätzt (vgl. Moed und Visser 2008). Der SSCI als Teil des Web of Science war jahrelang die einzige Quelle für die Zitationsanalyse. Die Unterkategorie Volkswirtschaftslehre im SSCI umfasst jedoch nur 247 Zeitschriften. In den letzten Jahren sind der kostenpflichtige Dienst Scopus und die kostenfreien Datenbanken Google Scholar und RePEc hinzugekommen. Letztere hat vor allem den Fokus auf den Wirtschaftswissenschaften, während die anderen beiden nahezu alle wissenschaftlichen Gebiete abdecken.<sup>1</sup>

Dieser Aufsatz greift den Beitrag »RePEc – eine unabhängige Plattform zur wirtschaftswissenschaftlichen Output-Messung« von Seiler und Wohlrabe (2010a) auf, in dem das RePEc-Netzwerk als unabhängige Plattform zur aktuellen Forschungsevaluation vorgestellt wurde. Das RePEc-Netzwerk ist eines von mehreren Möglichkeiten, Zitierungen zu extrahieren und, darauf aufbauend, Impaktfaktoren zu berechnen. Der vorliegende Artikel stellt, neben RePEc, den SSCI sowie

<sup>1</sup> Darüber hinaus gibt es, wie für die Wirtschaftswissenschaften, spezielle disziplinorientierte Datenbanken, beispielsweise Chemical Abstracts oder MathSciNet (vgl. Neuhaus und Daniel 2008 für weitere Informationen).

Google Scholar vor<sup>2</sup>, beschreibt deren Funktionsweisen und Besonderheiten und vergleicht die Netzwerke bezüglich ihrer Eigenschaften der Extraktion von Zitaten. Anschließend werden die drei Datenbanken hinsichtlich ihrer Fähigkeit zur Durchführung von Forschungsevaluation untersucht und Unterschiede bei der Berechnung von infometrischen Kennzahlen und Qualitätsmerkmalen, wie insbesondere dem Journal Impactfaktor (JIF), herausgearbeitet und kritisch bewertet. Die Analyse und Bewertung erfolgt vor allem im Hinblick auf die Verwendbarkeit in den Wirtschaftswissenschaften.

Wie bereits erwähnt, sind im Bereich Volkswirtschaft innerhalb des SSCI nur 247 Zeitschriften erfasst, für die ein Impactfaktor ausgewiesen wird. Dies ist sicherlich einer der Gründe, warum Alternativen zur Bewertung und/oder dem Ranking von Zeitschriften entwickelt wurden (vgl. beispielsweise Ritzenberger 2008; Schneider und Ursprung 2008; Schulze et al 2008; Combes und Linnermer 2010). Ritzenberger (2008) führt eine Vielzahl von Punkten an, die den Bias von verschiedenen Rankingverfahren erklären können. Der vorliegende Artikel befasst sich jedoch ausschließlich mit den Impactfaktoren und zeigt insbesondere auf, inwieweit Alternativen den Hauptnachteil der begrenzten Zeitschriftenauswahl für den Bereich Wirtschaftswissenschaften innerhalb des SSCI ausgleichen können, und empfiehlt, in Zukunft neue Rankings auf Basis neuer Datenbanken zu berechnen.

### Research Papers in Economics (RePEc)

Der bibliographische Service RePEc (Research Papers in Economics, [www.RePEc.org](http://www.RePEc.org)) wurde bereits in dem Beitrag Seiler und Wohlrabe (2010a) ausführlich vorgestellt. An dieser Stelle wird deshalb nur ein kurzer Überblick über das Netzwerk gegeben.

Das Ziel des RePEc-Netzwerkes ist es, wirtschaftswissenschaftliche Forschungsergebnisse, die in irgendeiner Form veröffentlicht wurden, zu sammeln und gleichzeitig mit diesen Informationen unterschiedliche Auswertungen durchzuführen. Eine Besonderheit von RePEc ist, dass es auf dem Mitmachprinzip beruht und nicht eine einzige Institution/Netzwerk alle relevanten Informationen zusammenträgt. So obliegt es den Verlagen, die Meta-Informationen ihrer Publikationen dem Netzwerk zur Verfügung stellen. Auf der anderen Seite müssen sich die Wissenschaftler bei RePEc registrieren und sich ihre Werke und Affiliations zuordnen. Ei-

ne weitere Besonderheit stellen die verschiedenen monatlichen Rankings dar, die RePEc auf Grundlage der bibliographischen Informationen erstellt. So werden Autoren- und Institutionenrankings sowohl auf Länder- als auch auf Kontinental- und Weltebene erstellt.

Das RePEc-Netzwerk sammelt auf zwei Arten Zitate. Zum einen besteht die Möglichkeit, dass die Verlage die Zitate als Meta-Informationen zur Verfügung stellen. Hier besteht nur teilweise Bereitschaft dies zu tun, da bestimmte Großverlage eigene Zitierungsdienste anbieten, wie z.B. Scopus von Elsevier. Zum anderen versucht RePEc, alle Artikel elektronisch automatisch auszuwerten, sofern sie frei zugänglich sind. Vor allem der letztere Punkt ist fehleranfällig.<sup>3</sup> Viele große Provider verlangen eine Gebühr für den Download eines Zeitschriftenartikels. Wenn sie frei zugänglich sind, bestehen noch größere Unsicherheiten bei der automatischen Extraktion der Zitate. Insbesondere Sonderzeichen, Umlaute und nicht-englische Artikel führen dazu, dass viele Zitate entweder fehlerhaft oder gar nicht ausgelesen werden können. Dieser Punkt ist sicherlich eine der größten Schwächen von RePEc (vgl. Seiler und Wohlrabe 2010b). Da es in den Wirtschaftswissenschaften durchaus üblich ist, vor der eigentlichen Veröffentlichung in einer Fachzeitschrift die Arbeit in einer Working-Paper-Reihe zu veröffentlichen, existieren oft mehrere Versionen einer Arbeit in RePEc. Wenn für eine Version die Zitate vorliegen und für die anderen nicht, wird hier die Annahme getroffen, dass die extrahierten Zitate auch in den anderen Versionen vorliegen. Sie werden deshalb für die jeweiligen Versionen übernommen. Diese Vorgehensweise ermöglicht es z.B., für nicht kostenfrei zugängliche elektronische Artikel die Zitate zu ergänzen. Ein potentieller Nachteil ist, dass ein Bias unbekannter Art entstehen kann. Dies ist dann der Fall, wenn nur einige bestimmte Arbeiten auch als Working Paper veröffentlicht werden.

RePEc extrahiert Zitate von allen indextierten wissenschaftlichen Arbeiten, d.h. auch auf Basis von Büchern und Working Papers. Ein Zitat als solches gilt erst als registriert, wenn sowohl der zitierende als auch der zitierte Artikel in RePEc gelistet sind. Dies ist ebenfalls ein Nachteil, da nicht alle (wirtschafts-)wissenschaftlichen Reihen in dem Netzwerk indextiert sind. Dies betrifft insbesondere interdisziplinäre Zeitschriften. Darüber hinaus besteht eine Limitierung, so dass teilweise nicht die komplette Historie einer Zeitschrift gelistet ist und Zitate aus älteren Ausgaben nicht berücksichtigt werden.

### Social Science Citation Index (SSCI)

Der SSCI ist eine Datenbank für Zeitschriftenliteratur der Sozialwissenschaften und, neben dem naturwissenschaftlichen

<sup>2</sup> Die Zitationsdatenbank Scopus hat im Grundsatz die gleiche Funktionsweise und Aufbau wie das Web of Science, nur liegt eine größere Zeitschriftenabdeckung vor. Es werden jedoch auf Basis der extrahierten Zitate keine Impactfaktoren oder Ähnliches ermittelt und veröffentlicht. Aufgrund des fehlenden Zugangs zu dem kostenpflichtigen Dienst wird hier auf eine nähere Erläuterung verzichtet.

<sup>3</sup> RePEc selber bezeichnet alle Informationen auf Basis von Zitaten noch als experimentell.

Science Citation Index (SCI) und dem geisteswissenschaftlichen Arts and Humanities Citation Index (A&HCI), Teil des Web of Science. Das Web of Science selbst ist eine multidisziplinäre Zitationsdatenbank und das Hauptprodukt des Institute for Scientific Information (ISI), welches zu dem Unternehmen Thomson Scientific gehört.

Im SSCI werden insgesamt 2 175 Zeitschriften ausgewertet (Stand: September 2010), was nur einen Teil der akademischen Zeitschriftenliteratur in den Sozialwissenschaften abdeckt. Die Kategorie Volkswirtschaftslehre (»Economics«) enthält sogar nur 247 Zeitschriften. Insgesamt umfasst das Web of Science mehr als 9 000 Zeitschriften, deren Zitate zueinander erfasst werden. Laut Thomson Scientific kommt der Auswahl der Journals eine große Bedeutung zu. Es sollen »Kernzeitschriften« der Wissenschaften ausgewertet und »Randbereiche« außer Acht gelassen werden (vgl. Stock und Stock 2003). Begründet wird dieses Vorgehen mit dem Bradfordschen Gesetz, wonach nur wenige Zeitschriften nach der Anzahl ihrer Artikel für eine Wissenschaftsdisziplin voll relevant sind. So zeigt Garfield (1979), dass 84% aller Zitate der im ISI vorhandenen Zeitschriften auf nur 2 000 Journals entfallen. Die Identifikation der Kernzeitschriften erfolgt Thomson Scientific zufolge nach Kriterien wie der Anzahl der Zitate einer Zeitschrift, der regelmäßigen und termingerechten Erscheinungsweise, dem Einhalten gewisser Konventionen (aussagekräftiger Zeitschriftentitel, aussagekräftiger Artikeltitel, vollständige bibliographische Angaben bei den Fuß- und Endnoten, Anschriften aller Autoren), dem Vorhandensein englischer Artikeltitel, Abstracts und Keywords (gerade bei nicht-englischsprachiger Literatur), sowie einem Peer-Review-Verfahren vor der Annahme angebotener Artikel (vgl. Stock 2001). Außerdem sei der Inhalt einer Zeitschrift von Bedeutung, da Zeitschriften, deren Bereich im ISI bereits vertreten ist, eine geringere Chance hätten, in den SSCI aufgenommen zu werden. Auch die internationale Ausrichtung der Zeitschrift spiele bei der Auswahl eine große Rolle (vgl. Stock, 2001). Es soll aber direkt an dieser Stelle kritisch bemerkt werden, dass Thomson Scientific keine zufriedenstellende Begründung für eine eventuelle Ablehnung bietet (vgl. Stock 2001).

Der SSCI bietet seinen Abonnenten die Möglichkeit, Fachartikel zu recherchieren und die Suchergebnisse zu speichern. Die Datensätze enthalten bibliographische Angaben, einen Abstract, Informationen zum Verlag, die Angabe der Referenzen und aller Zitate, so dass es möglich ist, sowohl zur zitierten als auch zur zitierenden Literatur zu navigieren. Die Volltexte können nur aufgerufen werden, wenn die Institution des Nutzers die entsprechende Zeitschrift abonniert hat. Interessant ist auch eine Suchoption, die es dem Nutzer ermöglicht, verwandte Artikel ausfindig zu machen. Hierbei wird nach der Übereinstimmung der Referenzen des Ausgangsartikels mit den anderen Dokumenten gesucht und eine Trefferliste sortiert nach der Anzahl gemeinsamer Zitate erstellt (vgl. Stock und Stock 2003).

Die Extraktion der Zitate erfolgt sowohl manuell als auch automatisch, d.h. es besteht eine nahezu vollständige Sicherheit, dass alle Zitate aus dem Zeitschriftenpool erfasst werden. Wie bereits oben erwähnt, ist dies auch der größte Nachteil. Es werden nur Zitate aus dem Zeitschriftenpool berücksichtigt, nicht aber andere Zeitschriften, Bücher oder Working Paper. Dieser Nachteil wiegt auch deshalb so schwer, da in den Sozialwissenschaften die wissenschaftliche Kommunikation nicht nur über Zeitschriften, sondern auch über Bücher erfolgt. Ähnlich dem RePEc-Netzwerk müssen sowohl die zitierende als auch die zitierte Zeitschrift innerhalb des Pools sein. Dies erlaubt eine eindeutige Zuordnung und vermeidet Doppelungen.

### Google Scholar (GS)

Google Scholar ist ein von Wissenschaftlern, Studierenden und Interessierten genutzter wissenschaftlicher Suchdienst, der sich durch seine Einfachheit, Suchgeschwindigkeit und Kostenfreiheit auszeichnet (vgl. Mayr und Walter 2005). Besonders attraktiv an Google Scholar ist, dass es seinen Nutzern neben einer einfachen Suchoption die Möglichkeit zur Durchführung einer kostenfreien Zitationsanalyse bietet. Dadurch wird Google Scholar zum unmittelbaren Konkurrenten diverse Fachdatenbanken.

Eine Besonderheit von Google Scholar ist, dass es bei der Suche nach wissenschaftlichen Dokumenten auf ein breites Spektrum an Organisationen, Verlagen und wissenschaftlichen Datenbanken zurückgreift. Dabei geht Google Scholar nach einem anderen Schema vor, als dies die herkömmliche Suchmaschine Google tut: es beschränkt sich nur auf nachweislich wissenschaftliche Informationen (vgl. Mayr und Walter 2005). Dieses wird durch die Kooperation mit Fachverlagen (z.B. Wiley Interscience oder Springer Verlag) und Wissenschaftsgesellschaften erreicht, deren Volltextbestände Google indexiert hat. Der Verlag Elsevier hat eine Indexierung durch Google abgelehnt. Der Grund ist sicherlich, wie bereits zuvor erwähnt, dass Elsevier mit Scopus eine eigene Datenbank anbietet. Da der Verlag auch viele wichtige wirtschaftswissenschaftliche Zeitschriften herausgibt, ist hier sicherlich eine größere Lücke zu konstatieren.

Neben diesen Volltextbeständen der Verlage beinhaltet Google Scholar aber auch Dokumente aus dem Open Access und Self-Archiving-Bereich – so auch RePEc (vgl. Mayr und Walter 2005). Der Recherchierende kann deshalb auf eine breite Menge an wissenschaftlichen Informationen zugreifen. Es sind nicht nur Artikel von angesehenen und international orientierten Journals indexiert, sondern auch Working Paper, Technical Reports, Dissertationen und Konferenzpapiere. Eine weitere Besonderheit von Google Scholar ist seine Interdisziplinarität, da es sich nicht nur auf ein bestimmtes Fachgebiet konzentriert und somit das Auffin-

**Box 1****Autonomous Citation Indexing (ACI)**

ACI ist ein automatisches Zitationsextraktionsverfahren, mit dem wissenschaftliche Dokumente gesucht und analysiert werden. Wurde ein Dokument identifiziert (z.B. durch einen Webcrawler oder durch eine Kooperation mit Verlagen), dann wird es in ein Textdokument umgewandelt und auf wissenschaftlichen Inhalt überprüft, indem nach einer Referenzliste gesucht wird. Nachdem bibliometrische Informationen, wie die URL des Dokuments, der Titel und Autor, ein Abstract (falls vorhanden), die Einleitung und die Referenzen entnommen wurden, werden die einzelnen Referenzen extrahiert und mittels speziellen Algorithmen Informationen wie Titel, Autor, Publikationsdatum, Zeitschrift und Jahrgang erkannt. Dabei sind die Algorithmen nach der »*Invariants-first*«-Methode konstruiert: Es werden diejenigen Zitationsbestandteile als nächstes identifiziert, die eine ähnliche Position und Zusammensetzung haben, wie die davor erkannten Zitationsbestandteile. Beispielsweise steht der Autor immer vor und das Journal nach dem Titel. Um die Autoren- und Journalnamen richtig zu identifizieren, werden Listen mit bereits bekannten Autoren und Journalen verwendet.

Ein wichtiger Bestandteil des ACI ist die Fähigkeit, verschiedene Arten von Zitaten des gleichen Artikels zu erkennen. Dazu werden alle Zitate zuerst normiert (Kleinschreibung, Entfernen von Bindestrichen und Fußnoten, Ausschreibung von geläufigen Abkürzungen, Entfernen von Phrasen, wie »forthcoming in«), nach ihrer Länge geordnet und danach mit einem *iterativen Verfahren* auf Übereinstimmung geprüft. Für jede Zitation wird die maximale Anzahl an Wörtern ermittelt, die sie mit einer anderen gemeinsam hat. Übersteigt diese Anzahl einen Schwellenwert, dann wird die neue Zitation demselben Artikel wie die vorhergehende Zitation zugeordnet. Im anderen Fall wird eine neue Gruppe für die Zitation erstellt. Neben diesem einfachen Algorithmus gibt es noch andere kompliziertere Verfahren, die bei Giles et al. (1999) beschrieben werden.

Da ACI vollkommen automatisiert ist und ohne menschlichen Eingriff arbeitet, ist es für Fehler anfällig. So kann es vorkommen, dass Zitate überschätzt werden, wenn verschiedene Versionen desselben Artikels online gefunden werden.

den von Verbindungen zwischen Forschungsdisziplinen ermöglicht. Jedoch kann es auch vorkommen, dass Dokumente gefunden werden, die nicht eindeutig wissenschaftlichen Quellen zuzuordnen sind, weshalb der Dienst mit Vorsicht zu genießen ist (vgl. Noruzi 2005). Weiterhin bleibt anzumerken, dass Google keine genauen Informationen darüber angibt, welche Quellen durch Google Scholar abgedeckt werden, welcher Zeitraum einbezogen ist oder welche Art von Dokumenten indexiert wurden (vgl. Mayr und Walter 2005).

Wie bereits erwähnt, besteht eine Stärke von Google Scholar darin, dass es den Nutzern eine Zitationsanalyse ermöglicht. Durch das Anklicken der »Zitiert-durch«-Option kann der Recherchierende nachvollziehen, durch wen und wie häufig ein bestimmter Artikel zitiert wurde. Die Ergebnisse dieser Suche werden nach bestimmten Kriterien geordnet, und es wird ein Relevanzranking erstellt. Das Hauptkriterium bei der Erstellung des Relevanzrankings ist die Anzahl der Zitierungen durch andere Dokumente. Dabei bedient sich Google Scholar einer automatischen Zitationsextraktion und -analyse (Autonomous Citation Indexing, ACI, vgl. Box 1), welche durch die Anwendung von Syntaxalgorithmen die Literaturangaben der analysierten Dokumente durchsucht und extrahiert. Damit ist es möglich, über die Referenzen analysierter Dokumente hinaus auch Literaturquellen nachzuweisen, die nicht auf den indexierten Webservern liegen (vgl. Mayr und Walter 2005). Es kann deshalb durchaus vorkommen, dass Google Scholar ein Ergebnis für einen nicht indexierten Artikel liefert. In diesem Fall wurde der Artikel über die Referenzen anderer Artikel gefunden. Für den Artikel besteht dann kein Hyperlink, er kann aber mit Google Web Search gesucht werden.

Problematisch an der »Zitiert-durch«-Option ist, dass Google Scholar tendenziell die Anzahl an Zitaten als zu hoch ausweist. Dieses liegt hauptsächlich an Problemen bei der Zitationsextraktion. Beispielsweise hat Google Scholar Probleme bei der Berücksichtigung gleicher Artikel, die auf verschiedenen Websites vorzufinden sind. Auch die Erkennung von bibliographischen Informationen wie Autorennamen, Publikationsdatum und die Zuordnung der Autoren zu Institutionen bereitet Probleme (vgl. Harzing und van der Wal 2009).

Ein Vorteil von Google Scholar ist seine Kostenfreiheit. Jedoch ist nur die Suche nach Fachartikeln kostenlos, da sich viele Inhalte auf Verlagsservern befinden, auf denen der Volltextabruf kostenpflichtig ist. Oft wird dem Recherchierenden aber ein Abstract zur Verfügung gestellt.

### Vergleich der Extraktionsverfahren

Jede der drei beschriebenen Datenbanken bzw. Internetplattformen hat Vor- und Nachteile. Ein Vergleich der Eigenschaften ist in Tabelle 1 angegeben.

Die umfassendste Extraktion, verbunden mit Exaktheit der Zitate, ist sicherlich innerhalb des SSCI gegeben, da viele Ergebnisse auch manuell überprüft werden.<sup>4</sup> Die Extraktion von Zitaten auf Basis automatisch elektronischer Verfahren, wie sie RePEc und Google Scholar teilweise einsetzen,

<sup>4</sup> Aber auch innerhalb des Web of Science gibt es Fehler bei der Extraktion der Zitate (vgl. beispielsweise die Fallstudie von Braun und Glänzel 1995). Inwieweit systematische Fehler vorliegen, lässt sich noch nicht beurteilen.

**Tab. 1**  
**Vergleich der Zitationsdatenbanken**

	Web of Science	Google Scholar	RePEc
Jahre der berücksichtigten Zeitschriften	Ab 1900 (Naturwissenschaften), ab 1956 (Sozialwissenschaften)	Nicht bekannt	Abhängig vom Verlag (teilweise ab 1896), teilweise nur unvollständig indiziert
Zeitraum für Zählung der Zitate	Ab 1900	Nicht bekannt	Abhängig von gelisteten Zeiträumen der Zeitschriften
Kostenpflichtig	Ja	Nein	Nein
Inhalte/Abdeckung	Mehr als 9 300 Zeitschriften (Natur, Sozial- und Kulturwissenschaften), 247 Zeitschriften im Bereich Volkswirtschaftslehre, keine Bücher, Buchkapitel <sup>a)</sup>	Nicht bekannt, enthält aber Bücher, Zeitschriften, Working Paper und Konferenzbände	Mehr als 1 100 Zeitschriften (vor allem Ökonomie und Statistik), knapp 3 000 Working-Paper-Reihen, Bücher und Buchkapitel
Infometrische Kennzahlen <sup>b)</sup>	Impaktfaktor (zwei und fünf Jahre), Immediacy Index, Halbwertszeit	Keine	Impaktfaktoren (einfacher, rekursiver, diskontiert, rekursiv diskontiert), <i>h</i> -Index

<sup>a)</sup> Einige wenige Konferenzbände und viel zitierte Buchreihen werden erfasst. – <sup>b)</sup> Von der Datenbank zur Verfügung gestellt.

Quelle: Zusammenstellung des ifo Instituts.

hat viele Fehlerquellen, wie z.B. Erkennung doppelter Dokumente, Matching von zitierenden und zitierten Artikel oder Schreibfehler. Dies kann entweder zu einer fehlerhaften, keiner oder mehrfachen Erfassung von Zitaten führen. Bei RePEc werden vor allem Zitate nicht erkannt, während bei Google Scholar alle genannten Effekte auftreten können.<sup>5</sup>

Größter Nachteil des SSCI ist natürlich, dass der Pool der Zeitschriften begrenzt ist, vor allem im Bereich Volkswirtschaftslehre mit nur 247 Zeitschriften. Darüber hinaus werden keine Bücher, Working Paper oder Buchkapitel erfasst. Dieses wiederum wird durch Google Scholar und das RePEc-Netzwerk gewährleistet. Aber auch in diesen Fällen gibt es Begrenzungen, bei RePEc durch die Anzahl der gelisteten wissenschaftlichen Reihen, die jedoch erheblich größer ist als beim SSCI. Bei Google Scholar ist nicht eindeutig geklärt, welche Inhalte abgedeckt werden und welche nicht. In jedem Fall ist ein größerer Überblick über die Anzahl der Zitate, zumindest im Rahmen der Wirtschaftswissenschaften, möglich. Sowohl für das RePEc-Netzwerk als auch für den SSCI gilt das Prinzip, dass der zitierende und der zitierte Artikel im Netzwerk indiziert werden müssen. Auf die Limitierung dieses Ansatzes wurde bereits verwiesen, er hat jedoch den Vorteil, dass Zitierungen eindeutig zugeordnet werden können.

Allen drei Plattformen ist gemein, dass für einzelne Artikel die Anzahl der entsprechenden Zitierungen ermittelt werden kann. Sie eignen sich also dafür, wissenschaftliche Recher-

che durchzuführen. SSCI und RePEc erlauben darüber hinaus, infometrische Kennzahlen, wie Impaktfaktoren, zu berechnen, was mit Google Scholar nicht unmittelbar möglich ist. Erwähnt werden sollte noch, dass der SSCI kostenpflichtig ist, während die anderen beiden Plattformen frei im Internet zugänglich sind.

Abschließend kann sicherlich festgestellt werden, dass trotz einiger Lücken bei der Erfassung der Zitate in RePE, dieses Netzwerk dem SSCI um einiges überlegen ist. Auch Google Scholar besitzt eine breitere Auswertungsbasis als der SSCI, jedoch können ohne erheblichen Zusatzaufwand keine infometrischen Kennzahlen ermittelt werden.

### Qualitätsanalyse auf Basis von Zitaten

Um die Qualität von wissenschaftlichen Artikeln und Zeitschriften zu erfassen, kann mit Zuhilfenahme der bibliographischen Informationen eine infometrische Analyse durchgeführt werden. Im Folgenden werden die infometrischen Kennzahlen vorgestellt, die von Thomson Scientific's SSCI und dem RePEc-Netzwerk direkt berechnet und dem Nutzer zur Verfügung gestellt werden. Mit Google Scholar allein lässt sich direkt keine infometrische Analyse ausführen. Jedoch können mit spezieller Software, die ein Auswertungstool für Google Scholar bereithält, Zitationsanalysen von Autoren und Zeitschriften vorgenommen werden. Zunächst stellen wir die bekanntesten Kennzahlen vor und zeigen, inwieweit diese bei den jeweiligen Netzwerken angewendet werden.

Die bekannteste infometrische Kennzahl ist der Impaktfaktor (IF) einer Zeitschrift, welcher auf Garfield (1955; 1972) zu-

<sup>5</sup> Google Scholar hat somit die Tendenz, zu viele Zitate anzuzeigen, wenn bestimmte Artikel in mehreren Reihen veröffentlicht worden sind. So ist es in den Wirtschaftswissenschaften immer häufiger zu beobachten, dass einzelne Artikel in mehreren Working-Paper-Reihen (wie z.B. NBER, CESifo oder IZA) erscheinen.

rückgeht. In seiner einfachsten Form berechnet er sich folgendermaßen:

$$IF = \frac{\text{Anzahl der Zitate in anderen Zeitschriften}}{\text{Anzahl der Artikel}}$$

Ein IF von z.B. 2 bedeutet, dass ein Artikel in einer Zeitschrift durchschnittlich zweimal zitiert wurde. Bei der Interpretation gilt es zu beachten, dass schon ein einzelner Artikel, der häufig zitiert wird, ausreicht, um einen hohen IF zu erhalten, während alle anderen erschienenen Artikel wenig oder gar nicht zitiert werden. In der praktischen Anwendung wird die Auszählung der Zitate und Artikel auf einen bestimmten Zeitraum eingegrenzt. Der Standardimpaktfaktor innerhalb des SSCI berücksichtigt Zitate der jeweils letzten beiden Jahre. Darüber hinaus gibt es unterschiedliche Erweiterungen und Modifikationen des IF, die teilweise noch erläutert werden. Das Konzept des (Standard-)IF innerhalb des Web of Science ist vielerlei Kritik ausgesetzt u.a.:

- Es gibt keine Normalisierungen der IF. Diese würden es erlauben, IF über verschiedene Forschungsgebiete zu vergleichen.
- Die Art der Zitierung wird nicht berücksichtigt, d.h. jede Zitierung unabhängig von der Zeitschrift erhält das gleiche Gewicht.
- Die Berücksichtigung von nur zwei Jahren für die Auszählung von Zitierungen ist zu gering.
- Es gibt einen Bias für Zeitschriften mit längeren Artikeln (z.B. Review Journals).

Für weitere Kritikpunkte und wie für diese korrigiert werden kann, siehe u.a. den Überblicksartikel von Glänzel und Moed (2002). Ein hoher IF kann auch auf die Tatsache zurückzuführen sein, dass nur wenige Artikel sehr häufig, während andere Artikel nie zitiert werden. Seglen (1994) zeigt, dass 15% der Artikel für 50% des IF einer Zeitschrift verantwortlich sind.<sup>6</sup> Trotzdem kann der IF immer noch als einflussreichstes infometrisches Kriterium für wissenschaftliche Zeitschriften angesehen werden.<sup>7</sup>

Auf Basis der Zitierungen wird auch der so genannte *h*-Index von Hirsch (2005) berechnet. Dieser ist wie folgt definiert: Ein Wissenschaftler hat einen Index von *h*, wenn von seinen *n* Artikeln *h* mindestens *h*-mal zitiert wurden. Die verbleibenden (*n*-*h*)-Artikel haben nicht mehr als *h*-Zitierungen. Dieser Index spiegelt damit sowohl die Qualität als auch die Breite der Forschungsergebnisse wider. Wissenschaftler mit wenigen, aber sehr oft zitierten Arbeiten haben ten-

denziell einen geringen *h*-Index. Ruane und Tol (2008) sowie Tol (2009) und Ellison (2010) geben Beispiele für den *h*-Index im Bereich der Volkswirtschaftslehre.

Der *h*-Index ist ein immer stärker beachteter Indikator zur Bewertung von Wissenschaftlern (vgl. Ellison 2010). Inzwischen ist dieser weiter entwickelt worden, um einige Nachteile auszugleichen. So ist z.B. eine Eigenschaft des *h*-Index, dass er viel zitierte Aufsätze nur unverhältnismäßig berücksichtigt (vgl. Clermont 2010). Darüber hinaus wird nicht für die Anzahl der Koautoren korrigiert. Im RePEc-Netzwerk hat beispielsweise der Harvard-Wissenschaftler Andrei Shleifer einen Indexwert von 54, einige seiner Artikel wurden aber mehrere Hundert Mal zitiert. Dieses kann sowohl als Vorteil gesehen werden, da dann stark zitierte Artikel (»one hit wonder«) nicht so stark gewichtet werden (vgl. Glänzel 2006), aber auch als Nachteil, da Highlights in der Publikationskarriere eines Wissenschaftlers damit auch weniger gewürdigt werden (vgl. Harzing und van der Wal 2008). Es kann somit auch passieren, dass Wissenschaftler zwar einen gleichen *h*-Indexwert aufweisen, die absolute Anzahl ihrer Zitate sich jedoch stark unterscheiden. Um diesem Problem Rechnung zu tragen, wurde von Egghe (2006) der *g*-Index entwickelt. Bei einer nach Anzahl der erhaltenen Zitate sortierten Publikationsliste ist der *g*-Index derjenige Rangplatz, bei dem *g* Publikationen in der Summe eine Zitationsmenge von mindestens *g*<sup>2</sup> erreichen (vgl. Egghe 2006). Artikel mit einer hohen Anzahl an Publikationen bekommen in der Indexberechnung ein höheres Gewicht. So besagt ein Wert von 32, dass ein Wissenschaftler 32 Artikel veröffentlicht hat, die gemeinsam mindestens 32<sup>2</sup> = 961 Zitate auf sich vereinen. Eine Verschärfung des *h*-Index ist der so genannte Wu-Index (vgl. Wu 2008). Dieser wird analog zum *h*-Index berechnet, jedoch benötigt man das 10*w*-fache an Zitaten, um einen Wert von *w* zu erreichen (für eine Verallgemeinerung des *h*-Indexes, welcher den klassischen *h*-Index, den *g*-Index und den Wu-Index als Spezialfall enthält, vgl. Ellison 2010).

#### Infometrische Analyse auf Grundlage des SSCI

Unter Zuhilfenahme der bibliographischen Daten aus dem SSCI lassen sich infometrische Kennzahlen, wie die Zitationsrate, der IF, der Immediacy Index und die Halbwertszeit, für die jeweiligen im SSCI gelisteten Journals berechnen. Diese Kennzahlen werden auf jährlicher Basis in den Journal Citation Reports (JCR) von Thomson Scientific veröffentlicht.<sup>8</sup>

Thomson Scientifics IF, der auch oft als klassischer IF bezeichnet wird, ist wahrscheinlich die populärste infometri-

<sup>6</sup> Frey und Rost (2010) weisen zu Recht darauf hin, dass aufgrund des IF nicht unbedingt auf die Qualität des einzelnen Artikels dieser Zeitschrift geschlossen werden kann.

<sup>7</sup> Dies ist unter anderem daran zu erkennen, dass nahezu alle Verlage den IF von ihren Zeitschriften auf den entsprechenden Websites angeben.

<sup>8</sup> Eine ausführliche Beschreibung der von Thomson Scientific berechneten infometrischen Kennzahlen findet man bei Stock (2001) sowie bei Garfield (2006).

sche Kennzahl. Berechnet wird er, indem die Anzahl der Zitate im Jahr  $t$  der Artikel aus den Jahren  $t - 1$  und  $t - 2$  durch die Anzahl der erschienenen Artikel in den Jahren  $t - 1$  und  $t - 2$  geteilt wird. Ein IF größer 1 bedeutet damit, dass mehr Zitate der Artikel aus den vergangenen zwei Jahren existieren, als Artikel in diesem Zeitraum publiziert wurden. Die SSCI-Plattform behandelt alle Zitate gleich. Problematisch ist dabei, dass Selbstzitationen von Journals berücksichtigt werden und dass das Verhältnis der Zitate zwischen einzelnen Zeitschriften unterschiedlicher Qualität unberücksichtigt bleibt.

Besonders aus Sicht der Wirtschaftswissenschaft scheint der geringe Beobachtungszeitraum problematisch und kann somit zu verzerrten Interpretationen des IF führen. Die Publikationsgeschwindigkeit beträgt in der Wirtschaftswissenschaft aufgrund des langen Begutachtungsprozesses oft über zwei Jahre (vgl. Ellison 2002). Ein Zitat einer bestimmten Zeitschrift kann dann nicht von dem IF erfasst werden, obwohl der Artikel in dem Zeitraum von zwei Jahren geschrieben wurde. Ein Beispiel zur Illustration ist in Box 2 gegeben. Eine direkte Folge des langsamen Publikationsprozesses ist, dass die IF aus dem wirtschaftswissenschaftlichen Bereich tendenziell niedriger sind als die aus den Naturwissenschaften. Aus Sicht der Wirtschaftswissenschaft sollten deshalb IF benutzt werden, deren Beobachtungszeitraum sich über mehr als zwei Jahre erstreckt.<sup>9</sup>

Ein weiteres Problem des kurzen Beobachtungszeitraums ist, dass aktuelle Zeitschriften, die direkt konsumiert werden, gegenüber den Journals, deren Zitate sich über einen längeren Zeitraum erstrecken, bevorzugt werden (vgl. Stock 2001). Um dieses Problem näher zu erläutern, wird eine konzeptionelle Basis für Alterungsprozesse wissenschaftlicher Zeitschriften benötigt. Glänzel und Schoepflin (1995) unterscheiden jeweils »kurze« und »lange« Perioden der »Reife« und des »Niedergangs« im Laufe des »Lebens« wissenschaftlicher Informationen und benennen vier Arten des Alterungsverhaltens von Journals:

1. kurze Reife, kurzer Niedergang: Informationen werden schnell rezipiert und schnell vergessen;
2. kurze Reife, langer Niedergang: Informationen werden schnell rezipiert und veralten nur langsam;
3. lange Reife, kurzer Niedergang: Informationen brauchen eine längere Zeit, um rezipiert zu werden, werden aber schnell wieder vergessen;
4. lange Reife, langer Niedergang: Informationen brauchen eine längere Zeit, um rezipiert zu werden, veralten aber nur langsam.

#### Box 2

##### Beispiel zur Zählung von Zitaten für einen Zweijahres-IF

Ein Wissenschaftler A schreibt im Jahr 2010 einen Artikel zu Prognosemethoden und zitiert einen Artikel des Wissenschaftlers B aus dem Jahr 2009, der in dem Journal *Econometric Theory* publiziert wurde. Der Artikel zu Prognosemethoden des Wissenschaftlers A erscheint aufgrund des langen Begutachtungsprozesses aber erst im Jahr 2012 in der Zeitschrift *Journal of Econometrics*.

Wird nun im Jahr 2012 der IF der Zeitschrift *Econometric Theory* innerhalb des SSCI ermittelt, dann findet die Zitation von B durch A keinen Eingang in diesen. Wie im Text beschrieben, werden zur Berechnung des IFs nur die Zitate von Artikeln aus den Jahren 2011 und 2010 berücksichtigt, nicht aber die Zitate des Artikels von Wissenschaftler A aus dem Jahr 2009.

Wie Stock (2001) bemerkt, kann der Zweijahres-IF nicht zwischen den Fällen unterscheiden. So fallen Differenzen zwischen den Fällen 1. und 2. nicht auf, und die Fälle 3. und 4. sind gar nicht zu erfassen, da die Reifephase wohl länger als zwei Jahre dauert. Auch hier wäre die Berechnung eines IFs mit längerem Beobachtungszeitraum sinnvoll.

Wie bereits erwähnt, ermittelt Thomson Scientific zwei weitere infometrische Kennzahlen: den Immediacy Index und die Halbwertzeit für die SSCI-Journals. Ersterer gibt an, wie schnell aktuelle Ergebnisse einer Zeitschrift von Wissenschaftlern rezipiert und weiterverarbeitet werden. Aufgrund der geringen Publikationsgeschwindigkeit in den Wirtschaftswissenschaften ist er kaum anwendbar. Der Immediacy Index scheint für die Wirtschaftswissenschaften weniger von Bedeutung zu sein und wird deshalb nicht ausführlich erläutert.

Die Halbwertzeit hingegen gibt Auskunft über die Relevanz einer Zeitschrift über mehrere Jahre hinweg und ist deshalb weitaus interessanter als der Immediacy Index. Sie gibt an, in welchen Zeitraum die Hälfte aller Zitate bzw. Referenzen fällt. Besteht ein Journal beispielsweise seit 1990 und vereinigt insgesamt 15 000 Zitate auf sich, von denen die letzten 7 500 Zitate auf die vergangenen fünf Jahre entfallen, dann hat das Journal eine Halbwertzeit von fünf Jahren. Geht man von dem Berichtsjahr 2010 aus, dann kommt man zu dem Schluss, dass das Journal in den letzten fünf Jahren genauso häufig zitiert wurde, wie in den ersten 15 Jahren seines Bestehens. Das Journal hat somit in den letzten fünf Jahren an Relevanz gewonnen.

#### Infometrische Analyse auf Grundlage von RePEc

Auch RePEc erstellt auf Basis der im Netzwerk verfügbaren Informationen eine Reihe infometrischer Kennzahlen, die die Grundlage für verschiedene Rankings für registrierte Autoren, Institutionen und Journal- und Working-Paper-Reihen sind. Anders als der SSCI errechnet RePEc dabei mehrere verschiedene IF. Der von RePEc berechnete *Simple-IF* unterscheidet sich von Thomsons IF dahinge-

<sup>9</sup> Thomson-Scientific weist auch einen Fünfjahres-IF aus.

hend, dass für jede Zeitschrift und Working-Paper-Reihe ein Wert berechnet wird. Außerdem unterliegt der *Simple-IF* keiner zeitlichen Begrenzung. Um ihn zu berechnen, wird die Anzahl der Zitate durch die Anzahl der insgesamt erschienenen Artikel geteilt. Aufgrund dieser Vorgehensweise werden zwar die Probleme des kurzen Beobachtungszeitraums umgangen, mit denen der Thomson Scientific IF behaftet ist, aber gleichzeitig wird ein Vergleich zwischen den Zeitschriften erschwert. Eine empirische Beobachtung ist, dass Zeitschriften mit wenigen Artikeln gegenüber Zeitschriften mit einer hohen Artikelanzahl tendenziell einen höheren IF aufweisen. Ein Beispiel ist in Tabelle 4 gegeben. Die CFS-Working Papers weisen einen IF von mehr als 10 auf, was sicherlich darauf zurückzuführen ist, dass im Oktober 2010 nur 169 Arbeiten gelistet waren, was im Vergleich zu den 3 158 der CESifo-Working-Paper-Reihe relativ wenig ist. Auch Zeitschriften, die in der Vergangenheit einige so genannte »Zitationsklassiker« publiziert haben, werden dadurch tendenziell einen höheren IF aufweisen.

**Tab. 2**  
**Vergleich der Anzahl der Zitate**

	SSCI	Google Scholar	RePEc
Daron Acemoglu, Simon Johnson und James A. Robinson (2001), »The Colonial Origins of Comparative Development: An Empirical Investigation«, <i>American Economic Review</i> 91(5), 1369–1401.	742	3 761	771
Rafael La Porta, Florencio Lopez-de-Silanes, Andrei Shleifer und Robert W. Vishny (1998), »Law and Finance«, <i>Journal of Political Economy</i> 106(6), 1113–1155.	1 499	5 212	1 118
Kahneman, Daniel und Tversky, Amos (1979), »Prospect Theory: An Analysis of Decision under Risk«, <i>Econometrica</i> 47(2), 263–291.	6 150	15 991	1 075
Ernst Fehr und Klaus M. Schmidt (1999), »A Theory Of Fairness, Competition, And Cooperation«, <i>Quarterly Journal of Economics</i> 114(3), 817–868.	909	3 411	636
Heckman, James, Ichimura, Hidehiko und Todd, Petra E (1997), »Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme«, <i>Review of Economic Studies</i> 64(4), 605–654.	397	42	307

Quelle: SSCI, Google Scholar, RePEc, aufgerufen am 13. Oktober 2010.

Zusätzlich zu diesem einfachen IF berechnet RePEc noch eine Reihe anderer IF:

- Damit nicht alle Zitate mit dem gleichen Gewicht in die Berechnung einfließen, d.h. Zitate aus Top-Zeitschriften nicht genauso bewertet werden wie Zitate aus Zeitschriften mit geringerer Qualität, wird der *rekursive IF* berechnet. Dabei ist die durchschnittliche Zitation auf einen Wert von 1 normiert. Dieser rekursive IF bemisst die Bedeutung der Zeitschriften relativ zueinander.
- Der *diskontierte IF* berücksichtigt das Alter der Zitation, indem er jede Zitation durch ihr Alter dividiert. So ergibt sich beispielsweise in Jahr 2010 für eine Zitation aus dem Jahr 2007 ein Wert von 0,25 (auch das aktuelle Jahr wird gezählt).
- Darüber hinaus gibt es noch den *diskontierten rekursiven IF*. Dieser beachtet sowohl die relative Bedeutung der Zeitschriften zueinander, als auch das Alter der Zitate.

Eine weitere infometrische Kennzahl, die RePEc angibt, ist der bereits vorgestellte *h-Index* von Hirsch (2005).

#### Infometrische Analyse auf Grundlage von Google Scholar

Wie bereits erwähnt, bietet Google Scholar keine direkten Möglichkeiten, infometrische Kennzahlen zu berechnen. Diesbezügliche Abhilfe schafft das kostenlose Auswertungstool *Publish or Perish (PoP)*<sup>10</sup>, mit dessen Hilfe eine Zita-

tionsanalyse von Autoren und Zeitschriften auf Grundlage von Google Scholar vorgenommen werden kann (vgl. Clermont 2010). Neben einfachen Kennzahlen zu der Anzahl von Dokumenten und Zitaten lassen sich mit *PoP* auch komplexere Indexwerte, wie der *h-Index* oder der *g-Index*, sowie Erweiterungen dieser Indices berechnen. IF können mit *PoP* jedoch nicht ermittelt werden.

#### Ein vergleichendes Beispiel

##### Zitate einzelner Artikel

Zunächst wird die Anzahl der erfassten Zitate für einzelne Artikel für alle drei Datenbanken verglichen. Zu diesem Zweck wurden die meistzitierten Artikel im SSCI der so genannten Top-5-Zeitschriften und die entsprechenden Zitate in Google Scholar und RePEc ermittelt.<sup>11</sup> Bis auf Kahneman und Tversky (1979) sind es Artikel, die in den letzten 15 Jahren veröffentlicht wurden. Dies ist zunächst nicht einleuchtend, da zu erwarten wäre, dass vor allem ältere Artikel mehr Zitate auf sich vereinigen. Ob die Erklärung eher in der zunehmenden Anzahl der wirtschaftswissenschaftlichen Zeitschriften, dem Zitierverhalten von Autoren oder der Qualität

<sup>10</sup> Die Software kann kostenlos heruntergeladen werden unter: <http://www.harzing.com/pop.htm>.

<sup>11</sup> Für weitere Vergleiche zwischen Google Scholar und Web of Science vgl. u.a. Norris und Oppenheim (2007) oder Bar-Illan et al. (2007).

der einzelnen Artikel geschuldet ist, kann in künftiger Forschung geklärt werden.

Es zeigt sich, dass im Vergleich zu SSCI und RePEc die ermittelten Zitate von Google Scholar in fast allen Fällen erheblich höher ausfallen. Die einzige Ausnahme ist der Artikel von Heckman et al. (1997), hier scheint Google Scholar die tatsächliche Anzahl der Zitate zu unterschätzen, denn die Anzahl von 42 Zitaten ist eindeutig zu niedrig. Die Höhe der Zitate zwischen SSCI und RePEc ist ungefähr gleich, mit der Ausnahme von Kahneman und Tversky (1979). Dies ist vor allem damit zu erklären, dass es sich um einen interdisziplinären Artikel handelt, der auch in sehr vielen anderen sozialwissenschaftlichen Zeitschriften zitiert wird. Abschließend ist zu bemerken, dass die RePEc-Zahlen mit der fortschreitenden zunehmenden Erfassung der Zitate noch steigen werden. Inwieweit sich diese in der absoluten Höhe den Google-Scholar-Zahlen angleichen, bleibt abzuwarten.

### Quantitative Beurteilung von Zeitschriften

In einem zweiten Schritt findet ein Vergleich infometrischer Kennzahlen für Zeitschriften auf Basis des SSCI und RePEc statt. In Tabelle 3 sind die Top-20-Zeitschriften gelistet, ermittelt auf Basis der höchsten IF im SSCI mit einem korrespondierenden IF in RePEc. Die Anzahl der Zitate deutet darauf hin, dass mit einer höheren Artikelzahl auch eine höhere Anzahl von Zitaten (absolut) erzielt wird (positive Korrelation). Dies gilt für beide Datenbanken. Wie bereits erwähnt, haben Zeitschriften mit einer geringen Artikelanzahl tendenziell einen höheren IF. In der Tat scheint das hier der Fall zu sein. Betrachtet man jedoch das komplette Sample von SSCI und RePEc, zeigt sich, dass die Korrelation positiv ist, d.h. Zeitschriften mit einer höheren Artikelanzahl besitzen tendenziell einen höheren IF. Ausnahmen sind jedoch Zeitschriften, die vor allem Überblicksartikel, wie das *Journal of Economic Literature*, oder sehr wenige Artikel veröffentlichen, wie z.B. das *Journal of Economic Growth*. Aufgrund der unterschiedlichen Zeit-

**Tab. 3**  
**Vergleich infometrischer Kennzahlen für Zeitschriften**

	SSCI					RePEc					Google Scholar
	Rang	Impaktfaktor (zwei Jahre)	Zitate 2009	Artikel	Impaktfaktor (fünf Jahre)	Rang	Impaktfaktor	Anzahl der Zitate	Artikel	h-Index	h-Index
Journal of Economic Literature	1	6,919	5 018	18	8,922	2	30,990	23 304	752	83	208
Quarterly Journal of Economics	2	5,647	13 985	43	8,171	1	33,431	68 365	2 045	124	213
Journal of Financial Economics	3	4,020	12 058	93	5,675	9	16,591	29 133	1 756	74	306
Econometrica	4	4,000	20 643	61	5,321	3	28,483	89 520	3 143	132	154
Journal of Economic Geography	5	3,937	1 146	35	4,705	17	2,906	988	340	16	60
Journal of Political Economy	6	3,841	16 350	31	6,924	8	17,342	92 535	5 336	144	245
Journal of Finance	7	3,764	18 039	78	6,536	11	11,384	46 809	4 112	91	297
Review of Environmental Economics and Policy	8	3,645	140	17	3,645	19	0,964	53	55	4	19
Journal of Economic Perspectives	9	3,557	5 649	41	5,38	7	17,770	23 011	1 295	73	87
Economic Geography	10	3,452	1 311	20	3,075	20	0,073	4	55	1	11
Experimental Economics	11	3,300	594	28	3,272	15	5,463	1 475	270	18	48
Journal of Economic Growth	12	3,083	1 075	12	4,967	4	27,794	5 781	208	39	86
Review of Economic Studies	13	2,904	7 068	50	3,926	6	18,915	37 072	1 960	82	176
Journal of Accounting and Economics	14	2,605	2 819	29	3,931	16	3,837	2 881	751	24	161
Journal of Environmental Economics and Management	15	2,581	3 177	49	2,967	14	6,622	9 608	1 451	37	126
Review of Economics and Statistics	16	2,555	7 319	59	4,044	13	9,315	31 391	3 370	69	129
American Economic Review	17	2,531	26 199	196	4,009	10	15,320	121 823	7 952	136	216
Ecological Economics	18	2,422	6 081	273	2,858	18	1,200	3 763	3 138	21	109
Economic Policy	19	2,375	891	16	3,211	5	19,174	3 528	184	31	77
Journal of International Economics	20	2,271	3 664	72	2,988	12	10,358	24 371	2 353	66	172

Quelle: SSCI, RePEc (September 2010).

**Tab. 4**  
**Vergleich infometrischer Kennzahlen für Working Paper in RePEc**

Rang	Working-Paper-Reihe	Artikel	Einfacher IF	Rekursiver IF	Diskontierter IF	<i>h</i> -Index
1	CFS Working Paper Series	169	10,716	1,121	3,66	22
2	CESifo Working Papers	3 158	3,788	0,248	1,15	39
3	IZA Discussion Papers	5 228	3,111	0,193	1,08	40
4	DIW Discussion Papers	1 053	2,040	0,089	0,69	19
5	ZEW Discussion Papers	1 241	1,971	0,081	0,61	21
6	Kiel Working Papers	744	1,411	0,067	0,49	15
7	Ifo Working Papers	92	1,272	0,056	0,50	5
8	Ruhr Economic Papers	201	0,403	0,030	0,23	4
9	IMK Working Papers	67	0,239	0,003	0,15	2
10	IWH Discussion Papers	193	0,192	0,003	0,07	3
	NBER Working Papers	16 445	14,374	1,091	3,5	167
	CEPR Working Papers	7 929	10,979	1,07	2,91	105

Quelle: RePEc (September 2010).

räume, auf die sich die IF beziehen, lassen sie sich zwischen SSCI und RePEc nicht direkt vergleichen. Betrachtet man jedoch die Reihenfolge, so zeigen sich einige Unterschiede. Die Zeitschriften *Journal of Economic Literature*, *Quarterly Journal of Economics* und *Econometrica* sind in beiden Rankings jeweils unter den Top 4. Die Zeitschrift *Economic Policy* ist im SSCI-Ranking auf Platz 19, während sie bei RePEc auf Platz 5 gelistet ist. Über die Ursache der verschiedenen Reihenfolgen lässt sich nur spekulieren. Einerseits ist es möglich, dass die fehlenden Zitierungen in RePEc zu einer systematischen Verzerrung führen. Andererseits kann dies auch ein Indiz dafür sein, dass das SSCI-Ranking verzerrt ist, da ein Großteil der Zitate von ökonomischen Zeitschriften nicht erfasst wird.

Tabelle 3 offenbart aber auch einige bereits erwähnte Schwächen des RePEc-Netzwerkes. So sind nur 55 Artikel der Zeitschrift *Economic Geography* vom Verlag gelistet (ab 2009). Die Zeitschrift wird jedoch schon seit 1925 herausgegeben. So ist die äußerst geringe Anzahl von Zitaten und der entsprechend niedrige IF zu erklären, welche die Bedeutung der Zeitschrift in keiner Weise im SSCI widerspiegelt. Die letzten beiden Spalten zeigen den *h*-Index für alle Zeitschriften jeweils auf Basis von Google Scholar und RePEc an.<sup>12</sup> Je höher der *h*-Index desto größer ist die Breitenwirkung der Zeitschrift, d.h. es ist eher unwahrscheinlich, dass ein hoher IF nur durch einzelne wenige Artikel entstanden ist. Insgesamt ist nahezu für alle Zeitschriften der *h*-Index, basierend auf Google Scholar, höher als bei RePEc. Dies ist vor allem der Tatsache geschuldet, dass Google Scholar im Moment noch mehr Zitate erfasst, und der bereits erwähnten zu kurzen erfassten Historie einiger Zeitschriften.

**Quantitative Beurteilung von Working-Paper-Reihen**

Abschließend illustriert Tabelle 4, dass der wissenschaftliche Einfluss auch von Working-Paper-Reihen erfasst werden kann.

<sup>12</sup> Die Zahlen für Google Scholar wurden mit der bereits erwähnten Software »Publish or Perish« ermittelt.

In den letzten Jahren ist es im Bereich der Wirtschaftswissenschaften mehr und mehr üblich, die eigene Forschung im Vorfeld der tatsächlichen Veröffentlichung in einer Fachzeitschrift in einer Working-Paper-Reihe zu veröffentlichen. Da durch die Möglichkeiten des Open Source und des Internets die Kosten einer Veröffentlichung erheblich gesunken sind, ist eine Vielzahl von Reihen entstanden.<sup>13</sup> So listet RePEc im Moment fast 3 000 verschiedene Reihen. Insofern stellen Impaktfaktoren eine Möglichkeit dar, in dieser großen Menge eine Qualitätseinschätzung vorzunehmen. In Tabelle 4 sind zehn deutsche Working-Paper-Reihen einschließlich ihrer verschiedenen infometrischen Maße dargestellt (sortiert nach dem einfachen IF). Bis auf den *h*-Index ist die CFS-(Center for Financial Studies)-Working-Paper-Reihe am einflussreichsten. Der hohe IF ist sicherlich auch darauf zurückzuführen, dass bisher nur 169 Papiere beim RePEc gelistet sind. Um die relative Qualität weltweit einzuordnen, sind die beiden bekanntesten Working-Paper-Reihen vom NBER und CEPR gelistet. Letztere sind im Vergleich in nahezu allen Kategorien den zehn deutschen Reihen überlegen.

**Abschließende Bemerkungen**

Die Einschätzung der Qualität von wissenschaftlicher Forschung ist sowohl für die Forschungsgemeinde als auch die Öffentlichkeit wichtig. Trotz der Kritik ist eine quantitative Beurteilung notwendig. In der Wissenschaft spielen die Zitate und die daraus abgeleiteten Impaktfaktoren eine wichtige Rolle, auch in den Wirtschaftswissenschaften. Dieser Artikel stellt drei Zitationsdatenbanken und ihre Funktionsweise vor, das Web of Science, welches den Social Science Citation Index enthält, Google Scholar und RePEc. Letztere hat einen speziellen Fokus auf die Wirtschaftswissenschaften.

<sup>13</sup> Die Möglichkeit in Working-Paper-Reihen zu veröffentlichen, unterscheidet sich von Reihe zu Reihe. So muss z.B. ein Autor Mitglied in den einzelnen Netzwerken sein, wie z.B. NBER, CESifo oder IZA, um ein Papier in der jeweiligen Reihe zu veröffentlichen.

ten. Google Scholar und RePEc haben den Vorteil, dass sie kostenlos im Internet verfügbar sind und eine Vielzahl von infometrischen Kennzahlen zur quantitativen Qualitätsbeurteilung bereitstellen. Darüber hinaus werden die Vor- und Nachteile von Impaktfaktoren, einer der wichtigsten Kennzahlen für wissenschaftliche Zeitschriften, herausgearbeitet. Mit Fokus auf die Wirtschaftswissenschaften ist sicherlich das RePEc-Netzwerk das am besten geeignete Forum zur Beurteilung von (quantitativer) Maßzahlen sowohl für Zeitschriften als auch Working-Paper-Reihen. Jedoch ist die tatsächliche Erfassung der Zitate noch unvollständig, so dass alle Kennzahlen noch als experimentell betrachtet werden sollten.

## Literatur

- Bar-Ilan, J., M. Levene und A. Lin (2007), »Some measures for comparing citation databases«, *Journal of Infometrics* 1, 26–34.
- Bommer, R. und H.W. Ursprung (1998), »Spieglein, Spieglein an der Wand – Eine publikationsanalytische Erfassung der Forschungsleistungen volkswirtschaftlicher Fachbereiche in Deutschland, Österreich und der Schweiz«, *Zeitschrift für Wirtschafts- und Sozialwissenschaften* 119, 1–28.
- Braun, T. und W. Glänzel (1995), »On a source of error in computing Journal Impact Factors«, *Chemical Intelligencer*, January, 31.
- Clermont, M. (2010), »Publish or Perish als Auswertungstool für Google Scholar – Aufbau und Auswertungsmöglichkeiten«, *Wirtschaftswissenschaftliches Studium* 39, 418–425.
- Combes, P.-P. und L. Linnemer (2010), »Inferring Missing Citations: A Quantitative Multi-Criteria Ranking of all Journals in Economics«, Working Paper.
- Coupe, T. (2003), »Revealed performance: Worldwide rankings of economists and economic departments, 1990–2000«, *Journal of the European Economic Association* 1, 1309–1345.
- Egghe, L. (2006), »Theory and practice of the g-Index«, *Scientometrics* 69, 131–152.
- Ellison, G. (2002), »The Slowdown of the Economics Publishing Process«, *Journal of Political Economy* 110, 947–993.
- Ellison, G. (2010), »How does the market use citation data? The Hirsch index in economics«, *CESifo Working Paper* No. 3188.
- Frey, B.S. und K. Rost (2010), »Do rankings reflect research quality?«, *Journal of Applied Economics* 13, 1–38.
- Garfield, E. (1955), »Citation indexes for science: New dimension in documentation through association of ideas«, *Science* 122, 108–111.
- Garfield, E. (1972), »Citation analysis as a tool in journal evaluation«, *Science* 178, 471–479.
- Garfield, E. (1979), *Citation Indexing – Its Theory and Application in Science, Technology and Humanities*, Wiley, New York u.a.
- Giles, C.L., K. Bollacker und S. Lawrence (1998), »CiteSeer: An Automatic Citation Indexing System«, *Digital Libraries 98: Third ACM Conf. on Digital Libraries*, ACM Press, New York, 89–98.
- Glänzel, W. (2006), »On the opportunities and limitations of the h-Index«, Working Paper, Leuven.
- Glänzel, W. und H.F. Moed (2002), »Journal impact measures in bibliometric research«, *Scientometrics* 53(2), 171–193.
- Harzing, A.W. und R. van der Wal (2008), »Google Scholar as a new source for citation analysis«, *Ethics in Science and Environmental Politics* 8, 62–71.
- Harzing, A.W. und R. van der Wal (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 & Technology* 60, 41–46.
- Hirsch, J.E. (2005), »An index to quantify an individual's scientific research output«, *Proceedings of the National Academy of Sciences* 102, 16569–16572.
- Mayr, P. und A.-K. Walter (2005), »Google Scholar – Wie tief gräbt diese Suchmaschine?«, Paper vorgetragen auf der 11. IuK-Jahrestagung: *In die Zukunft publizieren: Herausforderungen an das Publizieren und die Informationsversorgung in den Wissenschaften*, Bonn, Deutschland, 9–11. Mai 2005, erhältlich unter: [www.ib.hu-berlin.de/~mayr/arbeiten/Mayr\\_Walter05-preprint.pdf](http://www.ib.hu-berlin.de/~mayr/arbeiten/Mayr_Walter05-preprint.pdf) (aufgerufen am 28. September 2010).
- Moed, H. und M. Visser (2008), *Appraisal of citation data sources*, Centre for Science and Technology, Leiden University, Leiden.
- Nederhof, A. (2006), »Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A review«, *Scientometrics* 66, 81–100.
- Neuhaus, C. und H.-D. Daniel (2008), »Data sources for performing citation analysis: An overview«, *Journal of Documentation* 64, 193–210.
- Norris, M. und C. Oppenheim (2007), »Comparing alternatives to the Web of Science for coverage of the social sciences' literature«, *Journal of Infometrics* 1, 161–169.
- Noruzi, A. (2005), »Google Scholar: The new generation of citation indexes«, *Libri* 55, 170–180.
- Pislyakov, V. (2009), »Comparing two ›thermometers‹: Impact factors of 20 leading economic journals according to Journal Citation Reports and Scopus«, *Scientometrics* 79, 541–550.
- Ritzenberger, K. (2008), »A ranking of journals in economics and related fields«, *German Economic Review* 9, 402–430.
- Ruane, F. und R.S.J. Tol (2008), »Rational (Successive) h-indices: An Application to Economics in the Republic of Ireland«, *Scientometrics* 75, 395–405.
- Schneider, F. und H.W. Ursprung (2008), »The 2008 GEJ Journal-Ranking for the Economics Profession«, *German Economic Review* 9, 532–538.
- Schulze, G.C., S. Warning und C. Wiermann (2008), »Zeitschriftenrankings für die Wirtschaftswissenschaften – Konstruktion eines umfassenden Metaindexes«, *Perspektiven der Wirtschaftspolitik* 9, 286–305.
- Seglen, P.O. (1994), »Causal relationship between article citedness and journal impact«, *Journal of the American Society for Information Sciences* 45, 1–11.
- Seiler, Chr. und K. Wohlrabe (2010a), »RePEc – eine unabhängige Plattform zur wirtschaftswissenschaftlichen Output-Messung«, *ifo Schnelldienst* 63(7), 43–48.
- Seiler, Chr. und K. Wohlrabe (2010b), »Academic Rankings with RePEc: Some remarks«, *ifo Working Paper*, im Druck.
- Stock, M. und W.G. Stock (2003), »Wissenschaftliche Artikel, Patente und deren Zitationen: Der Wissenschaftsmarkt im Fokus«, *Password* (10), 30–37.
- Stock, W.G. (2001), »JCR on the Web. Journal Citation Reports: Ein Impact Factor für Bibliotheken, Verlage und Autoren?«, *Password* (5), 21–30.
- Tol, R.S.J. (2009), »The h-index and its Alternatives: An Application to the 100 Most Prolific Economists«, *Scientometrics* 80, 317–24.
- Ursprung, H.W. und M. Zimmer (2007), »Who is the ›Platz-Hirsch‹ of the German economic profession? A citation analysis«, *Jahrbücher für Nationalökonomie und Statistik* 227, 187–208.